THE
LONDON EDINBURGH AND DUBLIN
PHILOSOPHICAL MAGAZINE
AND
JOURNAL OF SCIENCE.

[SEVENTH SERIES.]

JANUARY 1931.

I. *The Theory and Performance of certain Types of Modern Acoustic Apparatus for reproducing Speech and Music.* *By* N. W. McLachlan, *D.Sc., M.I.E.E., F.Inst.P., and* G. A. V. Sowter, *B.Sc., AM.I.E.E.**

[Plates I.–III.]

Synopsis.

1. System with Elastic Constraint.
2. Discussion of Tabular Data relating to 1.
3. Natural Frequencies of Reed-driven Flexible Disk.
4. "Modes" of a Conical Diaphragm.
5. Data illustrating the Combination of Reed and Conical Diaphragm.
6. Necessity for an Elastic Member in a Diaphragm System.
7. Resonance of Coil-driven Diaphragm on its Surround.
8. Modes of a Moving Coil *per se*.
9. Linear Action in Diaphragms.
10. Accession to Inertia of Vibrating Diaphragm.
11. Amplitude of a Rigid Disk to Radiate a definite amount of Power.
12. Amplitude Limit imposed by Mechanical Stress in Reed.
13. Radial Velocity of Sound in a Flexible Disk.
14. Radial Velocity of Sound in a Conical Diaphragm.
15. Reproduction of Transients.
16. Graphical Illustration of Wave-form cited in 15.
17. Comparison between Theoretical and Actual Transients.
18. Causes of Distortion in a Moving Coil System.

* Communicated by the Authors.

Appendix.

Tables of Functions G and $K_1(z)$.

Table showing effect of Mutual Induction between Fixed and Moving Coils of a Reproducer.

ABSTRACT.

IN a former paper published in this Journal analyses were given of both coil- and reed-driven rigid disks as applied to the reproduction of sound. The object of the present contribution is to point a comparison between theory and practice, using these analyses. It is shown that both coil- and reed-driven disks, *where the whole moving system is quite rigid*, would be of little value in the reproduction of speech and music. The energy from the coil-driven disk decreases very rapidly above 1000 cycles (although good results are obtained on its axis), whereas the energy from the reed-driven disk is centered round the resonance frequency of the combination.

The case of a reed-driven *flexible* disk is analysed, and the results confirmed by experiment. It fails as a reproducer owing to the prominent (low damped) natural frequencies and the wide gaps between them where the radiation is small, due to the limited number of modes of the disk in the audible register.

In practice conical and reed-driven diaphragms give better results than those to be expected from the rigid disk theory. This is due mainly in the coil case to the elasticity of the general structure, and in both coil and reed drives to the "break-up" or modes of the diaphragm. The net result is an almost continuous succession of minor resonances and a tendency to uniformity in total radiation over a definite band of frequencies.

Various aspects of the problem are treated in detail, "modes" of the combinations being investigated and practice compared with theory. The problem of linear action in diaphragms is discussed, and it is shown that for a moderate output alien frequencies are not generated by the diaphragm, but may be caused by restriction of the axial motion at low frequencies due to an inelastic restoring force. In the reed drive even at moderate intensities, the excursion must be limited, not only for preservation of a linear current-deflexion relationship, but also to avoid undue bending stress in the reed at the support. Data are given to indicate the axial amplitude required to radiate a definite amount of power. It is shown that when the power is large, low frequencies cannot, in general, be radiated at full strength.

The radial velocity of sound has been measured in

aluminium and in steel disks, also near the periphery of a conical diaphragm. In the latter case it is considerably less than that of sound in air.

The reproduction of transients by a coil-driven rigid disk system is analysed, and the results compared with those obtained experimentally from coil- and reed-driven systems.

A summary of electrical, acoustical, and mechanical causes of distortion in a coil-driven reproducer is given, and the paper closes with an Appendix of Tables of various functions used in the calculation of data for design purposes.

---

1. *System with Elastic Constraint.*

THERE are two salient methods of driving large diaphragm * reproducing apparatus: (*a*) a circular coil situated in a radial magnetic field, (*b*) an elastic reed associated with a permanent magnet and fixed coils. The former method was treated at some length in a previous paper †, and it is now proposed to deal with the reed drive in detail. Also, various aspects of both types of drive will be discussed.

The general arrangement of reed-driven apparatus (assumed to be actuated by aid of a thermionic valve) is illustrated diagrammatically in fig. 1 *a*. In practice the reed, apart from the clamping nuts, usually has a natural frequency lying between 1500 and 3500 cycles.

The theory of this drive can be obtained from an analysis given in the former paper. When the reed is associated with a plane rigid circular disk, it was shown that the arrangement could be replaced by an equivalent electrical circuit. This circuit, with its components arranged *in series*, is illustrated in fig. 1 *b*. The driving mechanism can be replaced by an inductance L_0 of resistance R_0 (these being the values when the reed is clamped), in series with a motional inductance L_m or capacity C_m according to the frequency, and a motional or radiation resistance R_m. The two latter components are due to the motion of the reed in the magnetic field of the magnet, and their values vary with the frequency. The motional reactance X_m may be inductive or capacitive. The expressions for these components are :—

$$R_m = \frac{C^2\omega^2 B}{\omega^2 B + (k_1 - \omega^2 m)^2} = \frac{C^2 \cos^2\theta}{B}, \quad . \quad . \quad (1)$$

* Diaphragms whose projected radius exceeds 5 cm.

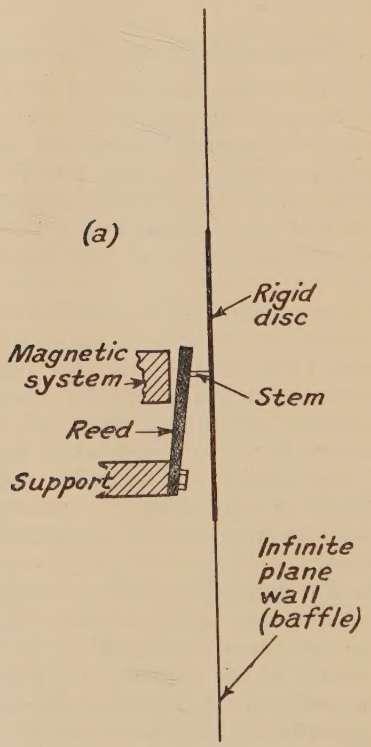
† Phil. Mag. vii. Suppl. June 1929, pp. 1011–1038.

where the acoustic power factor

$$\cos\theta = \frac{\omega B}{[\omega^2 B^2 + (k_1 - \omega^2 m)^2]^{\frac{1}{2}}},$$

$$L_m = \frac{C^2(k_1 - \omega^2 m)}{\omega^2 B^2 + (k_1 - \omega^2 m)^2}, \quad \ldots\ldots (2)$$

Fig. 1 *a*.



General arrangement of reed-driven rigid disk.

$$C_m = -\frac{[\omega^2 B^2 + (k_1 - \omega^2 m)^2]}{\omega^2 C^2 (k_1 - \omega^2 m)}, \quad \ldots\ldots (3)$$

$X_m = \omega L_m$ or $\frac{1}{\omega C_m}$ according to the frequency,

$\ldots$ (4)

where B = Acoustic pressure on disk per unit axial velocity,

C = Force on reed per unit current in coil,

k_1 = Effective coefficient of stiffness of reed,

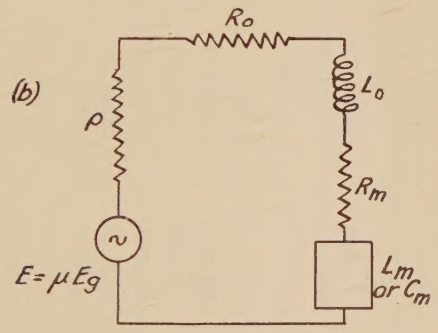
m { = Total effective mass of moving system,
{ = Natural effective mass + accession to inertia,

C_m = Motional capacity,

L_m = Motional inductance,

$\omega = 2\pi \times$ frequency.

Fig. 1*b*.



Equivalent electrical circuit (*series* arrangement of components) of fig. 1 *a*.

μ = Amplification factor of valve.

E_g = Voltage change on grid of valve (maximum).

E = Fictitious alternator corresponding to voltage change on anode of valve.

ρ = Internal A.C. resistance of thermionic valve.

R_0 = Effective resistance of driving mechanism when at rest.

L_0 = Effective inductance of driving mechanism when at rest.

R_m = Acoustic radiation resistance.

L_m = Motional inductance.

C_m = Motional capacity.

It is clear from (2) that, when $k_1 = \omega^2 m$ or $\omega = \left(\frac{k_1}{m}\right)^{\frac{1}{2}}$, the reactance is zero, *i. e.*, the system resonates. By aid of these formulæ and others given elsewhere * it is possible to calculate the performance of a reed-driven disk situated in a rigid plane of infinite extent.

* Phil. Mag. *loc. cit.*; Proc. Roy. Soc. A, cxxii. pp. 604–609 (1929).

TABLE I.—Mechanical Data pertaining to Reed-driven Rigid Disk.

Radius of disk (a)	21 cm.
Natural mass of reed * and disk (M)	56 gm.
Effective stiffness coefficient of reed (k_1)	$3{\cdot}6\times10^8$ dynes cm.$^{-1}$
Force on reed per unit current=e.m.f. induced in winding per unit velocity of reed (C)	2×10^8 dynes abs. unit^{-1}.
$k=\omega/c$ where c=velocity of sound in air	$3{\cdot}3\times10^4$ cm. sec.$^{-1}$

f. Frequency cycles per second.	$2ka=z$.	M_1. Accession to inertia. (gm.).	m. Effective mass (gm.).	B. Acoustic pressure on disk per unit velocity (dynes).	Cos θ. Acoustic power factor.	$CI_{r.m.s.}\cos\theta$. Acoustic pressure on disk (dynes).	$\ddot{x}_{r.m.s.}$ Axial acceleration of disk cm. sec.$^{-2}$	$P_{r.m.s.}$ Pressure on axis 10 metres from disk dynes cm.$^{-2}$
50	0·4	62·1	118·1	$2{\cdot}38\times10^3$	$2{\cdot}14\times10^{-3}$	$1{\cdot}56\times10^3$	$2{\cdot}06\times10^2$	$5{\cdot}9\times10^{-2}$
100	0·8	61·2	117·2	$8{\cdot}9\times10^3$	$1{\cdot}8\times10^{-2}$	$1{\cdot}29\times10^4$	$9{\cdot}1\times10^2$	$2{\cdot}61\times10^{-1}$
200	1·6	56	112	$3{\cdot}4\times10^4$	$2{\cdot}3\times10^{-1}$	$1{\cdot}53\times10^5$	$5{\cdot}66\times10^3$	1·62
305	**2·44**	**41·4**	**95·4**	$\mathbf{6{\cdot}9\times10^4}$	**1·0**	$\mathbf{5{\cdot}34\times10^5}$	$\mathbf{1{\cdot}48\times10^4}$	**4·25**
500	4·0	20·2	76·2	$12{\cdot}3\times10^4$	$6{\cdot}9\times10^{-1}$	$3{\cdot}61\times10^5$	$9{\cdot}22\times10^3$	2·64
1000	8·0	2·3	58·3	$11{\cdot}2\times10^4$	$3{\cdot}4\times10^{-1}$	$1{\cdot}33\times10^5$	$7{\cdot}45\times10^3$	2·13
2000	16·0	0·96	57	$11{\cdot}9\times10^4$	$1{\cdot}69\times10^{-1}$	$4{\cdot}33\times10^4$	$4{\cdot}56\times10^3$	1·31
4000	32·0	—	56	$11{\cdot}9\times10^4$	$8{\cdot}34\times10^{-2}$	$1{\cdot}17\times10^4$	$2{\cdot}47\times10^3$	$7{\cdot}07\times10^{-1}$
8000	64·0	—	56	$11{\cdot}9\times10^4$	$4{\cdot}3\times10^{-2}$	—	—	—

* The effective mass at the point of attachment to the disk is intended.

TABLE II.—Electrical Data pertaining to Reed-driven Rigid Disk.

Internal A.C. valve resistance $\rho = 2000$ ohms.

f. Frequency cycles per second.	$R_m = \frac{C \cos^2 \theta}{B}$. Calculated radiation resistance (ohms)	R_0. Measured resistance of coils with reed clamped (ohms).	$R_m + R_0 + \rho$. Total circuital resistance (ohms).	L_0. Measured inductance with reed clamped (henry).	L_m*. Calculated (henry).	$I_{r.m.s.}$ Current per 100 volt change on valve anode (amp.)	$R_m I^2_{r.m.s.}$ Energy radiated as sound (watts).
50	0·077	750	2750	0·67	0·115	$3·64 \times 10^{-2}$	$1·06 \times 10^{-3}$
100	1·46	800	2800	0·67	0·129	$3·57 \times 10^{-2}$	$1·86 \times 10^{-3}$
200	62·5	870	2930	0·67	0·208	$3·33 \times 10^{-2}$	$6·94 \times 10^{-2}$
305	**580**	**950**	**3530**	**0·66**	**0**	$\mathbf{2·67 \times 10^{-2}}$	$\mathbf{4·13 \times 10^{-1}}$
500	155	1100	3255	0·63	−0·051	$2·62 \times 10^{-2}$	$1·06 \times 10^{-1}$
1000	41·7	1550	3592	0·58	−0·018	$1·96 \times 10^{-2}$	$1·6 \times 10^{-2}$
2000	10·6	2530	4540	0·52	−0·0044	$1·28 \times 10^{-2}$	$1·7 \times 10^{-3}$
4000	2·57	4500	6500	0·50	−0·0011	$0·7 \times 10^{-2}$	$1·26 \times 10^{-4}$

* L_m is concomitant with the e.m.f. induced in the winding due to the motion of the reed in the magnetic field.

Numerical data pertaining to the mechanical and electrical aspects of this system are given in Tables I. and II. respectively.

The values of M_1, the accession to inertia *, and B, the acoustic pressure on the disk per unit velocity, were calculated by aid of the tabular values given in the Appendix. In calculating the value of B for frequencies of 1000 cycles upwards, the oscillation of the function G_1 was neglected. The size of the disk was chosen to tally with the projected area of a well-known commercial loud-speaker †, and its mass closely approximates to the mass of the cone used therein. The various electrical and mechanical parameters were obtained by measurement on the "Lion" movement. The value of k_1, the stiffness factor, was obtained by loading the reed with various masses, finding the corresponding resonance frequencies, and plotting $\frac{1}{\omega^2}$ against the added mass. The relationship was linear, k_1 being the slope of the line. The accurate measurement of C, the magnetic attraction on the reed per unit *direct current*, was accomplished by aid of an optical magnification system in which deflexions due to definite weights were compared with those due to the corresponding currents, or by using adequate current to bring the deflexion back to zero. It is desirable to use small deflexions in order to keep within the range where force and reed deflexion are linear. In fact, in this particular movement the plane of the reed is inclined at a small angle to that of the pole faces, so that the moment of the magnetic force causing deflexion will be proportional to the current over a restricted range as the reed moves *to or away from* the pole face. Care must be taken to avoid incorporating the deflexion of the housing in that of the reed.

Under operating conditions, owing to a phase displacement between the total alternating current and the flux, the value of C is different from that obtained with direct current. The alternating current in the winding can be resolved into a loss and a magnetizing component in quadrature. For accurate determination of C it is therefore necessary to multiply the total alternating current through the winding by $\sin\phi$, where ϕ is the angle of slope of the reactance-resistance curve of the movement at the frequency under consideration. The a.c. value of C is about 0·9 of its d.c. value.

* Phil. Mag. *loc. cit.*

† The Amplion "Lion," due to N.W.M.

It will be realized that the same argument applies to moving-coil apparatus.

The internal a.c. resistance of the valve was chosen as 2000 ohms, since this is a value which corresponds with the best practice for reed-driven apparatus. The reason for this is given in section 4*.

2. *Discussion of Data in Tables I. and II.*

A number of data in the Tables are of conspicuous interest. The resonance frequency of the system comprising the reed and disk oscillating as a rigid mass *in air* at N.T.P., is 305 cycles per second. *In vacuo*, where there is no accession to inertia, resonance occurs at 520 cycles. The natural frequency of the system in air is, of course, less than 305 cycles, owing to acoustic damping. At 305 cycles the acoustic power factor is unity, the radiation resistance, power output, and axial pressure have maximum values, whilst the motional inductance L_m is zero †, *i. e.*, it is at the change-over point from inductance to capacity. The inductance of the movement with the reed in action is $L_0 + L_m$, the sum of the static and motional inductances. The latter is created by the e.m.f. induced in the winding due to the motion of the reed causing a variation in the air-gap reluctance. It rises to a maximum value just before resonance, at which point it passes through zero, thence to a negative minimum, and thereafter approaches zero once again. The negative values imply a phase relationship which corresponds to a leading current, *i. e.*, a capacity effect. These variations in L_m are superposed on the static inductance curve, a condition which is clearly portrayed by the calculated values plotted in fig. 1 *c* and the actual measurements plotted in figs. 5 and 6. It should be observed that the values in the latter two figures are $L_0 + L_m$.

Owing to the neutralization of inertia by means of the reed compliance, the acoustic power factor, the radiation resistance, power output, and axial pressure *at resonance* are considerably greater than those for a coil-driven rigid disk at any frequency from zero upwards. A glance at fig. 2 will show that the output from the device, even on the axis, will be extremely unpleasant, since the bulk of the energy

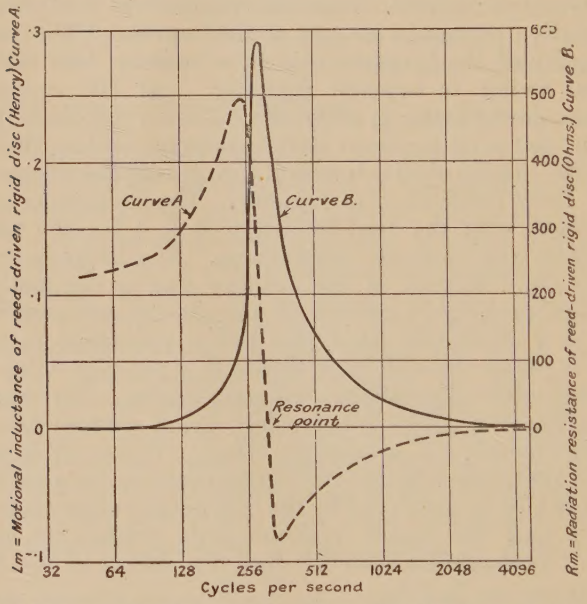
* See also 'Wireless World and Radio Review,' March 20th, 1929, pp. 295–301.

† It should be observed that L_m is not the inductance of the instrument whilst in action, but the additional inductance due to the vibration of the reed in the magnetic field of the magnet.

is concentrated over a comparatively narrow frequency band. Moreover, a rigid diaphragm system driven by an elastic reed would be of no value as a reproducer of speech and music.

In practice, however, reproducers of the reed type, although deficient in the lower musical register, give much better results than those to be inferred from the preceding data. A typical output curve taken on the axis of a practical reproducer is portrayed in fig. 7. Comparison between figs. 2 and 7 illustrates the enormous difference between a

Fig. 1 *c*.



Calculated curves, showing variation in motional inductance (Curve A) and in radiation resistance (Curve B) of reed movement driving rigid disk 21 cm. radius.

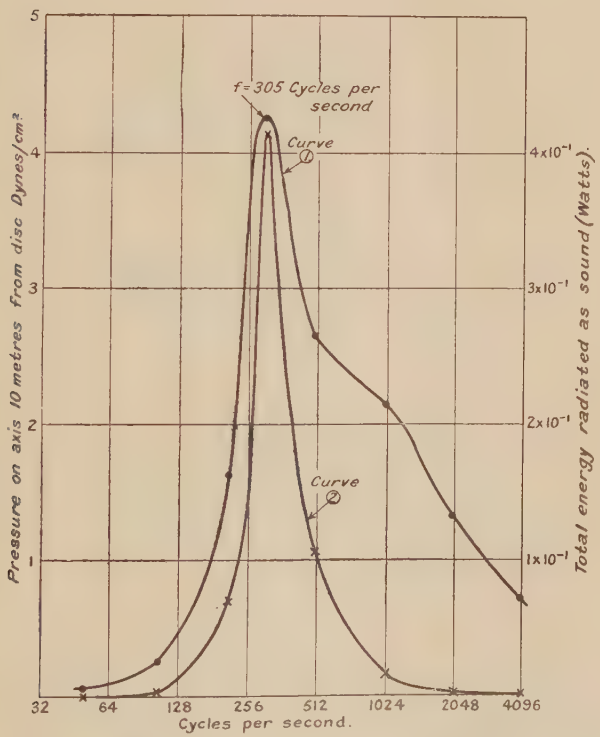
hypothetical and an actual reproducer. The reason for this is simply that the diaphragm does not move as a rigid structure, a physical phenomenon which has been invaluable in the evolution of large diaphragm reproducers for speech and music.

3. *Natural Frequencies of Reed-driven Flexible Disk.*

The study of a reed-driven conical diaphragm is best approached by taking a case amenable to analytical treatment.

It is proposed, therefore, to discuss the behaviour of a flat circular disk driven by an electromagnetically operated reed. A flexible circular disk exhibits various modes according to the conditions imposed at the edge, at the centre, or on the flat surface. In the present case, where the driving member is firmly clamped to a small area at the centre and the edge is free, we shall confine our attention to

Fig 2.



Curve 1.—Air-pressure on axis 10 metres from reed-driven rigid disk at various frequencies. (Last column of Table I.)

Curve 2.—Energy radiated as sound from reed-driven rigid disk. (Last column of Table II.)

circular modes. The appearance of radial modes depends upon certain conditions, *e.g.*, lack of homogeneity, asymmetry or eccentricity of drive on disk, etc. Such modes were obtained with free edge thin sheet steel disks, but not with an aluminium disk three times as thick.

Warren * has deduced formulæ for determining the apparent mass of a circular disk *in vacuo* under different conditions of constraint. The apparent mass of a disk having a *free edge* is given by the expression

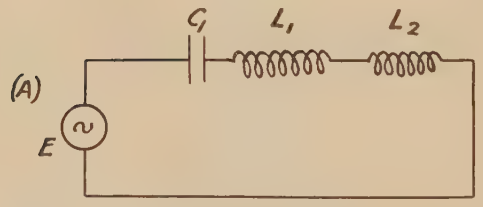
$$M_a = \frac{4\pi q}{h^2}\left\{\frac{\frac{B_1}{A_1}}{1+0{\cdot}11593\frac{B_1}{A_1}+\frac{C_1}{A_1}}\right\}, \quad . \quad . \quad . \quad (5)$$

where

$$\frac{B_1}{A_1}=\frac{a_2c_1-a_1c_2}{c_2(b_1+d_1)-c_1(b_2+d_2)},$$

$$\frac{C_1}{A_1}=-\frac{a_2(b_1+d_1)-a_1(b_2+d_2)}{c_2(b_1+d_1)-c_1(b_2+d_2)},$$

Fig. 3.



(A) Electrical analogue of reed-driven rigid disk.
E = Fictitious alternator corresponding to voltage change on anode of valve.
C_1 = Compliance of reed = 1/force per unit deflexion.
L_1 = Mass of reed.
L_2 = Mass of disk.

$a_1 = J_1(ha)$, $\quad a_2 = \frac{2(\sigma-1)}{\sigma ha}.J_1(ha)-J_0(ha)$,

$b_1 = Y_1(ha)$,

$c_1 = I_1(ha)$, $\quad b_2 = \frac{2(\sigma-1)}{\sigma ha}.Y_1(ha)-Y_0(ha)$,

$d_1 = -K_1(ha)$,

$c_2 = I_0(ha)$, $\quad a$ = radius of disk,

$d_2 = K_0(ha)$, $\quad h^4 = bg$,

σ = Poisson's ratio expressed >1, $\quad b = 2\pi\omega^2 q$,

* Phil. Mag. ix. p. 881 (May 1930). In order to avoid confusion with symbols used in the previous paper and also herein we have had to depart from those employed by Mr. Warren.

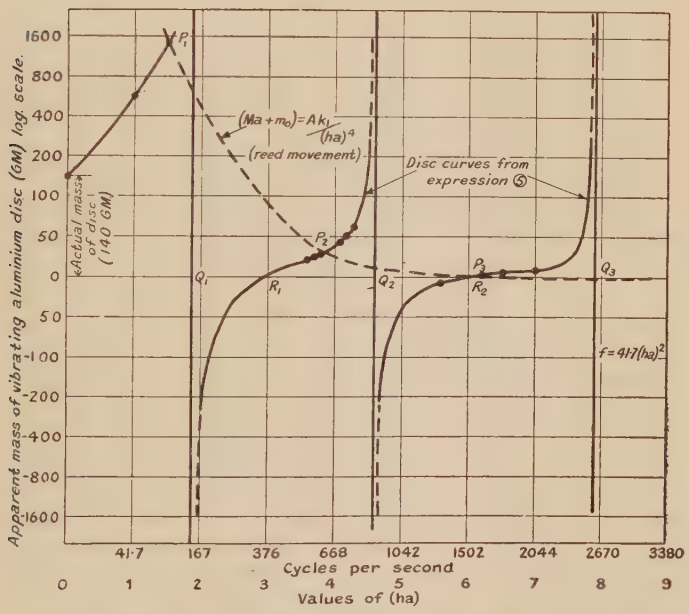
E = Youngs' Modulus of Elasticity, $\quad g = \frac{6}{\pi t^3} \cdot \frac{\sigma^2 - 1}{\sigma} \cdot \frac{1}{E}$,

t = thickness of disk, $\quad q$ = mass per unit area of disk.

The values of J_0, J_1, I_0, I_1, K_0, K_1 were obtained from Gray and Mathews' book on Bessel functions, whilst those of Y_0, Y_1 were taken from B. A. Report, p. 33 (1915).

When $A_1 + 0{\cdot}11593\,B_1 + C_1 = 0$, the apparent mass is infinite, and this condition corresponds to fixation of the centre.

Fig. 4.



Curves showing apparent mass of aluminium disk 10 cm. rad. 0·165 cm. thick, for various values of ha. Also dynamic relationship for vibrating reed loaded with varying mass. The relationship between frequency and ha for this particular case is $f = 41{\cdot}7\,(ha)^2$ or $\omega^2 = 1615\,(ha)^4$. At $R_1\,R_2$ the mass $= 0$ and $P_1\,P_2\,P_3$ are intersections.

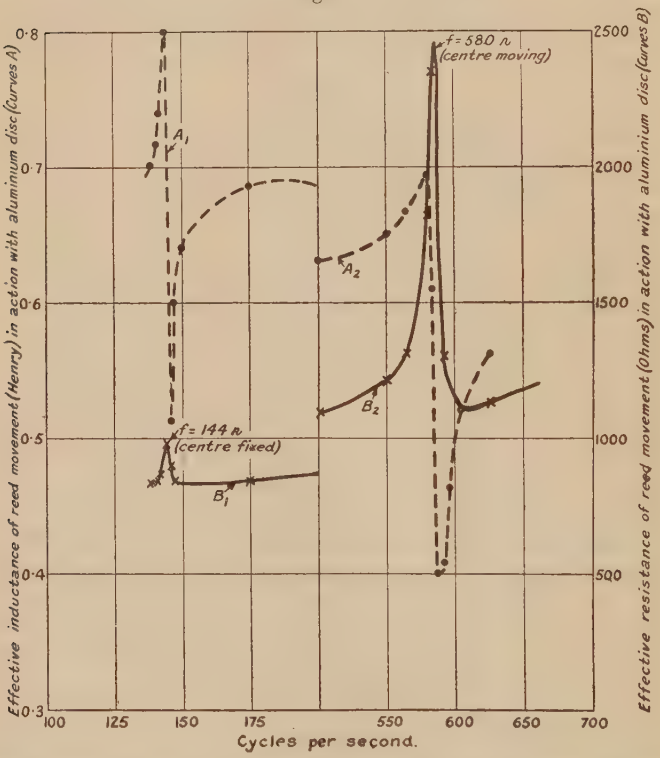
Evanescence of the apparent mass is concomitant with one of the symmetrical "free" modes of the disk (concentric circles as nodes), and for this condition $B_1 = 0$.

The values of M_a for a definite range of ha have been plotted in fig. 4 (full line curves) in connexion with an aluminium disk 10 cm. radius and 0·165 cm. thick. The

fixed and free frequencies occur respectively when $M_a = \infty$ and zero. The negative values of M_a can be regarded as arising from a change of phase between the driving and inertia forces associated with the disk.

There are two sets of natural frequencies to be considered : (*a*) those which are due to the combination of the reed and

Fig. 5.



Curves showing variation in total inductance, L_0+L_m, and total resistance, R_0+R_m, of reed movement driving aluminium disk 10 cm. rad. 0·165 cm. thick.

$f = 144 \sim$ = centre fixed, giving first mode of disk.
$f = 580 \sim$ = centre moving, giving first mode of combination.

The resistance variation at 580 ~ is nine times that at 144 ~. Compare with fig. 1 *c*.

disk, in which case the centre of the disk moves appreciably, (*b*) those due to the disk alone energized by the reed and vibrating with the centre substantially fixed, *i.e.*, a node.

Actually, there is a vibration of the centre under the stimulus of the reed to overcome frictional and radiation loss. In case (*a*) we have to satisfy the dynamic condition in expression (5) together with that for the reed, namely $\omega^2(M_a + m_0) = k_1$, where M_a is the apparent mass of the disk, m_0 is the effective mass of the reed and its attachment to the disk, and k_1 is the effective coefficient of restitution of the reed, the arrangement then being regarded as that of a simple loaded coil spring. The simultaneous solutions to (5) and $\omega^2(M_a + m_0) = k_1$ can be obtained by transforming the latter to

$$(M_a + m_0) = \frac{Ak_1}{(ha)^4},$$

where $A = 1/1615$, and plotting this relationship, using the M_a, *ha* axes of fig. 4 (dotted curves). The coordinates of the points of intersection (extending with *ha* to infinity, but finite in number within the audible range) yield the necessary data from which the natural frequencies for case (*a*) can be obtained. It should be observed that beyond a certain value of *ha* the value of M_a is negative, which means that the reed is virtually stiffened by the disk or that the phase of the motion is such as to neutralize part of the effective mass of the reed and the connecting link. Obviously the greatest negative value cannot exceed m_0.

The modes of oscillation, the corresponding frequencies, and the apparent masses obtained by experiment and from fig. 4 are set forth in Table III.

The modes were delineated experimentally in the well-known manner by sprinkling sand on the vibrating disk. The first mode of the *combination*, viz., at 90 cycles, does not yield a sand figure, since the whole arrangement moves. Difficulty was experienced in ascertaining the experimental value of frequency with certainty, since several resonances could be detected by ear, whilst no appreciable change in electrical impedance occurred. The second frequency of the combination determined empirically is 7 per cent. less than that calculated from expression (5), and this is due to the following:—(1) No allowance is made in the calculation for the added mass due to "accession to inertia"*. This would reduce the theoretical value. (2) E and σ were not measured, but merely extracted from physical tables.

The agreement at the third frequency is remarkably close.

The solutions to the case of fixed centre are those where the apparent mass $M_a = \pm\infty$. The corresponding values of

* The theory is for vibration *in vacuo*.

ha occur where the asymptotes are indicated in fig. 4. The calculated values of *ha* are 1·9, 4·7, 7·85, etc., corresponding to the frequencies 151, 920, and 2570, etc. The sound energy and the amplitude of vibration at the lowest frequency are quite appreciable, but the higher frequencies are difficult to detect, since the energy is small. Owing to this and to the whole structure vibrating slightly, due to the centre not being quite motionless, the sand figures (nodes) were not very clearly defined. The motion can be stimulated by gripping the centre of the disk over an appreciable area, say 1 cm. radius. The results obtained thus are shown in

TABLE III.

Combination Modes of Free Edge Reed-driven Circular Aluminium Disk.

Radius = 10 cm., thickness = 0·165 cm., mass = 140 gm.

Mode of oscillation.	*ha*.	M_a. Apparent mass of disk (gm.).	*f*. Frequency of mode. (cycles per second).	
			Theory (fig. 4).	Experiment.
No node. Disk moving throughout. Convex and concave alternately.....	1·5	1500 (theory).	90	Difficult to ascertain accurately.
Circular node. (radius 7·5 cm.).....	3·95	27 (expt.).	620	580
(Two circular nodes. radii 3·75, 8·25 cm.)	6·24	0·3 (expt.).	1630	1610

Table IV., from which once again the agreement between theory and practice is close. The lowest calculated frequency is, of course, too large, owing to the accession to inertia not being contemplated in the theory.

It is obvious that each of the natural frequencies of the system will be accompanied by variations in the electrical coefficients of the winding. Such variations can be measured in the well-known manner by the aid of a bridge. Measurements on the reed-driven aluminium disk are portrayed graphically in fig. 5.

If the centre were absolutely fixed at 144 cycles, there would obviously be no change in impedance. There is a large variation in inductance, but the variation in effective resistance is small compared with that at 580 cycles centre moving

—about $\frac{1}{9}$,—which indicates that the resonances with fixed centre are of less consequence aurally than those of the *combination* of disk and reed.

4. *Modes of a Conical Diaphragm.*

From a long series of experiments it has been found that a certain class of paper, when formed into a cone of appropriate radius and apex angle, makes a suitable diaphragm for reproducers *. The velocity of propagation of energy along generators of the cone decreases from the apex outwards. Its value depends upon the mechanical properties of the

TABLE IV.

Modes of Free Edge Reed-driven Aluminium Disk with Fixed Centre.

Mode.	ka.	f. Frequency of mode (cycles per second).	
		Theory.	Experiment.
Centre substantially at rest. Sand collects there.....	1·9	151	144
Centre substantially at rest. One circular node	4·7	920	900
Centre substantially at rest. Two circular nodes	7·85	2570	2600

paper and upon the geometrical dimensions of the cone. In general, as shown in section 14, it is less than the velocity of sound in air. This being so, the diaphragm cannot be regarded to move as a whole over the range of audible frequencies, unless it is really minute. Moreover, in practice the diaphragm exhibits the phenomenon of "breaking-up" or vibrating in such a manner that portions on either side of a node move in opposite directions. This introduces interference and results in a partial neutralization of the radiation from the diaphragm as a whole. The neutralization is due to two causes : (1) positive waves from one vibrating area and negative waves from another reaching points in space simultaneously; (2) circulation waves between oppositely vibrating areas due to the acoustic pressure difference †.

* See British Patent no. 288, 713.

† The same argument applies to the flexible disks discussed in section 3.

The frequencies at which a diaphragm breaks up depend upon its mechanical construction and geometrical form. The "shape" of the diaphragm when driven by an alternating force can be seen by aid of one or more neon tubes energized from a supply having a frequency nearly one half that supplied to the diaphragm. The amplitude of vibration at a peripheral antinode on a free edge cone can be quite large—of the order of 2 cm.—for moderate driving forces, whilst the aerial eddy

TABLE V.

Showing Radial Modes of a Free Edge Cone driven by a Circular Coil.

Radius of coil=2·5 cm.; projected radius of diaphragm=12·2 cm.; apicial angle (plane)=90°; thickness of paper=0·014 cm.

Frequency.	Number of radial nodes.	Frequency from curve.	Number of radial nodes.	Frequency ratio to fundamental.
36·5	4*	37	4	1
43	6	50	6	1·35
52	8			
60	10			
75	8	69	8	1·86
95	10	92	10	2·48
118	12	120	12	3·24
150	14	150	14	4·06
190	16	190	16	5·13
225	18	230	18	6·22

* The class of nodal figure is akin to that of the cylindrical ring cited in section 8, the first mode being characterized by four nodes.

currents at the edge are adequate to extinguish a lighted match. Moreover, at low frequencies the nodes can be explored roughly by this means or by a gas flame passed round the edge. By increasing the rigidity at the edge of the diaphragm, *e. g.*, bending it over and glueing to an annular rubber support, the modes occur at higher frequencies. In fact, at low frequencies the diaphragm behaves almost as if the edge were quite rigid, and when viewed by suitable intermittent light the portion between the edge and the apex is seen to exhibit radial modes.

The behaviour of a number of different types of diaphragm has been examined. The results obtained with a free edge conical diaphragm of thin paper driven annularly by a moving coil are given in Table V.

The coil was centred in the air gap of the magnet by a flexible paper former, and the cone supported near its periphery with a vertical silk thread. Owing to the flexibility of the diaphragm (thin paper with free edge) the first radial mode occurred at a very low frequency, and all the radial modes given in the Table were readily detected.

The cone was constructed in the well-known manner with a seam having an overlap 0·65 cm. wide, the two faces being treated with seccotine. This obviously destroyed the symmetry of the diaphragm and introduced an irregularity regarding the disposition of the third and fourth modes *. These modes were obtained at two different frequencies, but at the two lower frequencies, viz., 52 and 60 cycles, the spacing of the nodes was irregular. Moreover, the frequencies in column 3 have been obtained from a mean curve drawn through the experimental points.

This diaphragm was then inserted in a baffle 6 feet square, having a central hole 12·4 cm. radius, and tested on broadcasting from a high quality radio receiver of known performance. There was a conspicuous absence of the bass register and a relatively small output as compared with the same size of diaphragm of *thicker* paper stiffened and mounted at the edge on a flexible support. Clearly the reduction in the bass register is due to the break-up of the diaphragm occurring at too low a frequency by reason of (1) free edge, (2) too thin paper.

By supporting a cone and introducing a constraint on a circle between the apex and periphery, thereby leaving the edge free, the outer portion exhibits modes of large amplitude, whilst the inner portion moves as a whole (at low frequencies).

Clearly the acoustic radiation characteristic of a diaphragm depends upon the manner in which it moves at various frequencies, this being controlled by the mechanical construction. For example, to retain the lower frequencies it is essential that the diaphragm shall not be *too* large nor its edge free. A large diaphragm, even with constrained edge, or a small diaphragm with a free edge, means a low break-up frequency, resulting in partial neutralization of the radiation due to oppositely vibrating portions.

* Radial modes do recur at higher frequencies when circular modes appear. The effect, however, is not thought to be due to this.

There are two sets of modes to be considered :

(1) When the diaphragm is energized by sound radiation from another diaphragm and can execute forced or free oscillations without mechanical constraint imposed by a driving mechanism.

(2) When the diaphragm is driven centrally or annularly and executes either forced or free oscillations.

In our particular problem we are concerned with case (2). The diaphragm of the reproducer under consideration has a projected working radius of 21 cm., and is reinforced at the apex to a projected radius of 4·6 cm. It is damped at the periphery by cotton wool packed firmly on each side, the whole being mounted in a circular ring of vee section. Owing to the limitation in amplitude of the reed, we were unable to ascertain the fundamental break-up frequency of the combination. By aid of a neon tube it was seen that at frequencies in the neighbourhood of 100 cycles the reinforced centre moved as a whole, whilst the remainder of the diaphragm exhibited diametral modes. So far as the lower frequencies are concerned, therefore, the cone can be considered to be driven as a frustum (annular drive) with a constrained and damped outer edge. The fundamental break-up frequency is augmented by the edge condition, but occurs at a very low frequency.

Owing to this phenomenon the apparent mass of the diaphragm varies with the frequency. In fact, the reed can be regarded as being loaded with an amount (mass) of diaphragm which changes with rise in frequency. At different frequencies there are resonances due to the interaction of the reed and the diaphragm. There are also the diaphragm modes with fixed centre, although, judging from the disk experiments, such modes are not prominent acoustically. In addition, there is a certain degree of damping due to diaphragm transmission losses and to the radiation of sound. It is found that so far as the reed is concerned, the effective mass of the diaphragm is zero in the neighbourhood of 2000 * cycles ; also that the diaphragm contributes nothing to the radiation above a certain frequency (10,000 cycles in our case), *i. e.*, there is no effect on the sound output when the diaphragm is removed. By combining a reed whose natural frequency (including conical fixing nuts) is about 1600 to 2500 cycles with a diaphragm of suitable properties and geometrical configuration the resulting system

* It may be zero at a lower frequency, but this was not tested.

has no objectionable resonance to a steady input. The system can be represented electrically by an equivalent circuit of the form shown in fig. 1 *b*. In this case, however, the reactance X_m will be essentially complex and one which varies in structure with the frequency. An approximate electrical analogue (not the electrical equivalent) of the system, using a rigid disk, is illustrated in fig. 3.

The lateral motion of sound waves from the neighbourhood of the apex of the diaphragm on its concave side is conserved in a degree owing to the solid angle of the cone being less than 2π, and this is concomitant with an increase in energy. The reverse occurs on the convex side. Furthermore, the air pressure due to vibrations at the apex reaches the mouth of the cone before that from the periphery owing to the low velocity of propagation in the cone itself (see section 14). This will be evinced more particularly in reproducing a transient than a steady sound (see figs. 18, 19, 20). In the former case we are actually concerned with time relative to the initiation of the transient, whereas in the latter case it is only phase which matters.

5. *Practical Data illustrating the Combination of Reed and Conical Diaphragm.*

To illustrate these statements concretely a series of measurements of the effective inductance and effective resistance of a "Lion" loud speaker are reproduced in fig. 6 *. It is clear that when the reed is unrestricted there are numerous "resonances" or changes of reactance (variation of X_m of fig. 1 *b* throughout the frequency range indicated). Whilst the reed is clamped the system behaves as an ordinary iron-cored inductance. The difference between the inductance values with the reed at rest and in motion is equivalent to $X_m/\omega = L_m$, whilst the corresponding resistance value gives R_m the radiation resistance plus the diaphragm losses and additional iron losses due to the motion of the reed. The two latter can be found by taking additional readings *in vacuo* to eliminate the resistance component due purely to radiation.

Acoustic pressure measurements on the axis of a typical reed-driven conical diaphragm in free air at a distance of five diameters or more indicate a curve which rises from the lower frequencies up to a peak at 2000 cycles, this being the resonance frequency of the reed, stem, and conical nuts (as a

* The reed movement used for these tests was slightly different from that cited in Table II.

whole) used for attachment to the diaphragm (see fig. 7). Beyond 2000 cycles the pressure falls, but there is another rise at 5000 cycles due to an auxiliary resonance in the reed system. This is useful and improves the upper register. The output is definitely attenuated at the lower frequencies, a result one would expect from theoretical considerations.

Fig. 6.

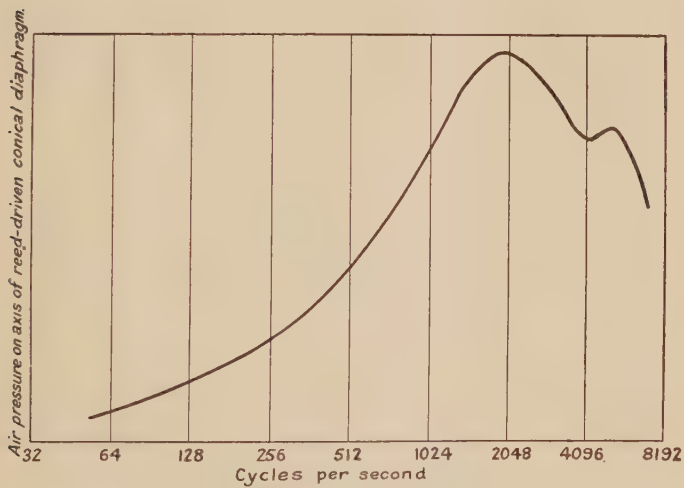
Curves showing variation in effective inductance (dotted curve) and resistance (full line curve) of reed movement driving conical diaphragm 21 cm. effective radius.

If the low frequency output is increased beyond a certain value, the amplitude of the reed—which increases with decrease in frequency—is excessive, the relationship between amplitude and current ceases to be linear and the reed is unduly stressed (see section 12). This is portrayed by the

record of fig. 8 (Pl. I.). Some movements of the reed type exhibit this effect in a more pronounced degree.

Although the axial pressure increases up to 2000 cycles the result in practice is decidedly better than one would anticipate from the curve of fig. 7. This is due to the focussing effect (interference), which gradually increases with the frequency *, Moreover, the total energy output does not increase with frequency, as shown by the pressure curve of fig. 7. Under ordinary listening conditions in an average room, where reflexion occurs from its six faces, it has been found by I. Wolff † that, provided the reproducer and the

Fig. 7.



Axial air pressure of reed-driven conical diaphragm.

listener are suitably situated, it is only the total energy radiated which matters. Thus, although the axial pressure may show a large increase at the upper frequencies, owing to focussing, the energy radiated may be no greater than at the middle frequencies. The result in a suitable room can be quite pleasing.

The curve of fig. 7 is an average one, since there are fluctuations owing to the various resonances caused by the diaphragm modes and to interference. With this particular

* Proc. Roy. Soc. *loc. cit.*
† Proc. I. R. E. xvi. pp. 1729–1741 (Dec. 1928).

reproducer the output falls away rapidly below 150 cycles per second.

It will be seen from column 7, Table II., that the current through the coils decreases with rise in frequency. In fact, at 50 cycles its value is five times that at 4000 cycles. This current variation is essential in order to offset the enhanced response as the natural frequency of the reed and cone nuts (2000 cycles) is approached. To secure a large low frequency current it is imperative to use a power valve of small internal a.c. resistance; also, to avoid reducing the upper register, the inductance of the movement must not be too high. The effective ohmic resistance of the movement should, of course, be as small as possible. Thus the design is a compromise between (*a*) natural frequency of reed, (*b*) geometrical configuration of diaphragm, (*c*) physical properties of diaphragm, (*d*) acoustical properties of diaphragm, (*e*) inductance and effective resistance of reed movement.

It is of considerable practical importance to point a comparison between the acoustic behaviour of a flexible disk and a conical diaphragm. As a reproducer of speech or music the former fails, owing to the prominent and widely spaced resonances which are unquestionably indicated in fig. 5. Clearly the acoustic output between the resonance peaks is small. No serious fluctuations occur between the main peaks in the curves, since the axial velocity and the driving force are well out of phase. The case of the paper cone is exhibited in fig. 6, where there are a large number of minor modes following each other rapidly, whilst apart from 116 cycles no really serious resonance occurs, due largely to the successive modes of break-up of the diaphragm and to transmission loss therein. Here the axial velocity and the driving force are much more frequently in closer phase relationship.

Moreover, the success of the paper cone as a reproducer over a certain part of the acoustic register is due to its lack of rigidity, to its many modes which occur at small frequency intervals, and to the damping of the material. In fact the modes change almost progressively with the frequency. Metal disks of the dimensions cited herein fail, since the modes are widely spaced, the damping small, and the mass considerable. To introduce a progressive series of modes at close frequency intervals without undue prominence of any particular mode the metal must be thin to reduce mass, stretched to an adequate degree of tension, and suitably damped.

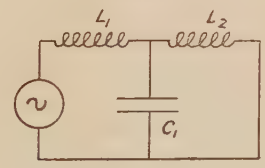
6. *Necessity for an Elastic Member in a Large Diaphragm System.*

In a former paper in this Journal it was shown that the energy output from a coil-driven rigid disk is extremely small above 1000 cycles—in fact this frequency can be regarded as a cut-off point. In practice, however, where the coil drives a conical diaphragm, it is sometimes found that the output above 1000 cycles is too powerful.

A conical paper diaphragm 0·02 to 0·025 cm. thick, apicial angle 90°, and peripheral radius 12·5 cm., mounted on a rubber surround in the usual manner, has its first radial mode at some frequency below 200 cycles per second. Above this frequency the radiation will tend to decrease, due to enhanced interference and to diaphragm transmission losses, unless the phenomenon of resonance occurs.

By a series of measurements under different conditions we discovered some years ago that the upper register in a moving

Fig. 9.



(*a*) Electrical analogue of coil-driven rigid disk having elastic member between coil and disk.

$L_1 \equiv$ Mass of coil.
$L_2 \equiv$ Mass of disk.
$C_1 \equiv$ Compliance of elastic member between coil and disk. 1/Force per unit extension.

coil reproducer is due to the resonating system formed by the coil and the diaphragm, the whole acting as a complex elastic structure *. In general there is more than one resonance. The position and magnitude of the resonances largely determine the merits of the reproducer, provided the remainder of the design is quite normal. The problem can be visualized more readily by considering the simple case of a coil-driven rigid disk with an elastic member between the coil and disk. The electrical analogue of this is illustrated in fig. 9. It is

* 'Wireless World and Radio Review,' pp. 497-499 (Oct. 10, 1928) and pp. 539–542 (Oct. 17, 1928).

of interest to indicate that the essential difference between the elastic member in reed- and coil-driven rigid disks is shown by the series and parallel connexions of C_1 in figs. 3 and 9.

The lowest resonance frequency of the combination is

$$f = \frac{1}{2\pi}[L_1C_1]^{-\frac{1}{2}}, \text{ } i.e., \text{ when } L_2 = \infty,$$

or the mass of the diaphragm is infinite.

In practice the coil and effective diaphragm masses are comparable and the resonance frequency is higher than the foregoing, since L_1 and L_2 are in parallel and give a value less than L_1.

For this case

$$f = \frac{1}{2\pi}\left[\frac{L_1L_2C_1}{L_1 + L_2}\right]^{-\frac{1}{2}}.$$

Apart from the modes of the coil *per se*—as set forth in detail in section 8—there are three major possibilities of resonance :—(1) An axial oscillation of the coil and the diaphragm, the neck of the coil between the winding and the diaphragm supplying the elastic member. (2) An axial oscillation of the coil and the neck (for attachment to the diaphragm) moving as a whole on the diaphragm. The latter, which can for the purpose of illustration be regarded as a disk of varying section, supplies the elastic member. The effective mass will vary, as in the case of the flexible disks discussed in section 3. (3) A combination of (1) and (2).

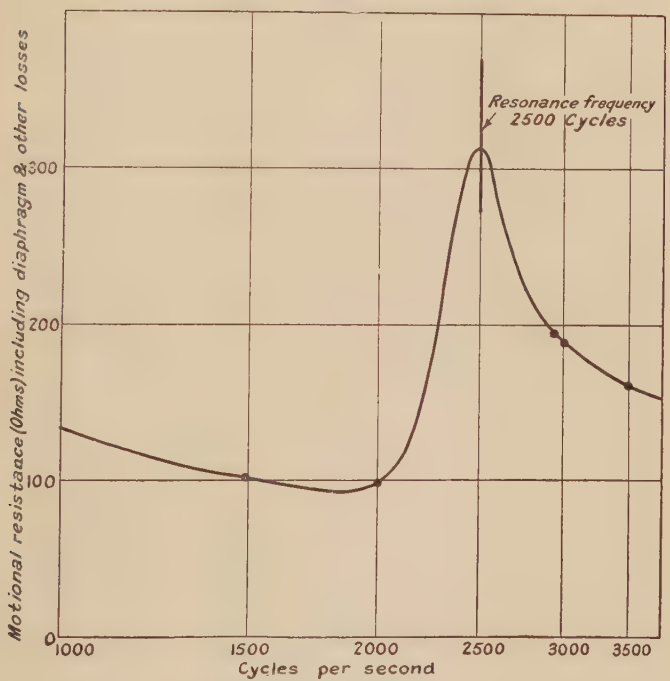
A discussion of the research work on the mechanical behaviour of diaphragms is too protracted to be included here. It is hoped, however, to present it in detail at a later date.

As a matter of interest a curve showing the difference in the effective resistance of the coil with the magnetic field off and on is given in fig. 10. This shows a decided resonance at 2500 cycles, due to the elasticity of the diaphragm structure, *i. e.*, case (2) above.

At a demonstration we witnessed recently a circular coil was used to drive a flat reinforced diaphragm of balsa wood. This instrument suffered in a marked degree from upper frequency resonances of the above nature, and the reproduction of speech and music was definitely distorted and somewhat distressing. The description of these diaphragms in the 'Journal of Scientific Instruments,' p. 51 (February 1930) would lead one to believe that they are a distinct

advance on the usual type of conical paper diaphragm. This, however, is not the case, and the data and correspondence published elsewhere * definitely establishes the fact that the upper register is much too powerful. Whether this can be rectified remains to be seen.

Fig. 10.



Curve illustrating resonance due to elasticity of diaphragm.

7. *Resonance of Coil-driven Diaphragm on Surround and of the latter* per se.

When the coil is constrained to move axially only there are two major low-frequency resonances † in the coil-driven

* 'Wireless World and Radio Review,' April 9th, p. 374; May 14th, p. 520; May 28th, p. 569; June 11th, p. 633; June 25th, pp. 673, 682 (all 1930).

† In practice there is a "wobble" frequency due to (1) the c. of g. of the diaphragm not being in the plane of the suspension, (2) the centre of magnetic force on the coil not passing through the c. of g. of the diaphragm, (3) lack of uniformity of the radial tension of the surround, (4) geometrical and mechanical asymmetry in general.

diaphragm : (*a*) the coil and diaphragm moving as a whole on the elastic surround at its periphery, which is used as a support, (*b*) the surround moving *per se* between the diaphragm and outer framework.

It is clear that in the absence of a magnetic field the frequency of the diaphragm on its surround is given by

$$f_0 = \frac{1}{2\pi}\left[\frac{k_0}{m}\right]^{\frac{1}{2}}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

where

k_0 = coefficient of restitution of the surround,

$m = \mathrm{M} + \mathrm{M}_1$ = effective mass of moving system, M_1 being the accession to inertia.

In practical reproducers the diaphragm is supported by some form of elastic or pseudo-elastic material. The natural frequency of the diaphragm on the support, without the damping effect of the magnetic field, should be well below audibility. When the support is stiff the resonance may occur at, say, 80 cycles, and this yields very objectionable reproduction of both speech and music. When a low-resistance moving coil is used—30 ohms or less—in combination with a transformer the magnetic field exerts a considerable damping on the coil, since the circuit is closed through the secondary winding of the transformer. With a weak support constraint axially, the motion of the coil is aperiodic *, *but the resonance to steady alternating currents still exists.* Where the axial constraint is strong, *e. g.*, natural $f = 80$ cycles, the motion is oscillatory. If the diaphragm is supported at the periphery on three threads it can be moved from side to side (wobble oscillation) by a puff of air, but will not move axially. This not only shows that the damping exerted axially is very large, but that it is essential that the supporting device should be able to eliminate sideways wobble, especially when reproducing very low frequencies, *e. g.*, pedal organ, where the axial amplitude is large. Additional information on this subject can be obtained by consultation of previous papers †.

The approximate frequency of the surround, considered after the manner of a vibrating string, is given by

$$f_1 = \left[\frac{1}{2}\frac{\mathrm{T}_1}{m_1 l}\right]^{\frac{1}{2}}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

* No support oscillation is visible in fig. 20.

† 'Wireless World and Radio Review,' pp. 154–157 (Aug. 8th, 1928); pp. 497–499 (Oct. 10th), pp. 539–542 (Oct. 17th, 1928); pp. 729–732 (Nov. 28th, 1928); pp. 243–245 (5th March, 1930).

where

T_1 = radial tension of surround,

m_1 = effective mass of surround, including accession to inertia,

l = radial depth of surround.

Referring to fig. 11, we have

$$T_1\theta = F = k_0 x,$$

or $$T_1 = \frac{k_0 x}{\theta}, \text{ since } x = l\theta.$$

Fig. 11.

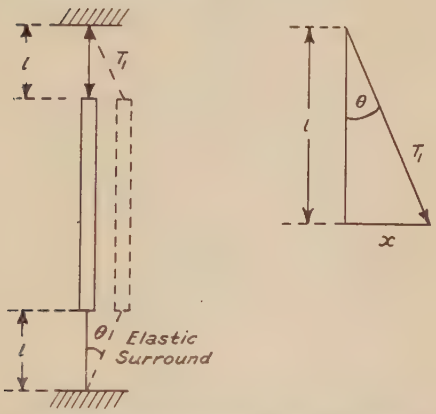


Diagram illustrating rigid circular disk with elastic surround.

Whence, substituting in expression (7), we get

$$f_1 = \frac{1}{2}\left[\frac{k_0}{m_1}\right]^{\frac{1}{2}}, \quad \ldots \ldots \quad (8)$$

and the ratio of the frequency of the surround to that of the diaphragm is

$$\frac{f_1}{f_0} = \pi \left[\frac{m}{m_1}\right]^{\frac{1}{2}}. \quad \ldots \ldots \quad (9)$$

In general $m > m_1$, so that the frequency of the surround may lie well within the audible register. We have assumed that the diaphragm mass is much greater than that of the surround, *i.e.*, the diaphragm has been treated as a rigid support. This is approximately correct, since m/m_1 will not be less than 8/1. Assuming the frequency of the diaphragm f_0 to be 15 cycles per second, that of the surround

will be nearly 130 cycles, which is well within the audible register. This shows the necessity for having f_0 as low as possible. We have found by measurement that a serious resonance can occur due to the surround. Its frequency varies with the radial tension of the surround, and this in turn alters with temperature and humidity. Values of f_1 from 50 to 200 ~ have been found, and the effect is readily seen on impulse records of the form shown in fig. 20.

8. *Modes of a Moving Coil* per se.

The simplest method of visualizing the modes of a coil by itself is to consider it as a ring of solid metal. With varnished and baked coils of fine enamelled wire (46 or 47 S.W.G.) this is approximately true. When the ring is quite free there are six forms of vibration, as follows :—

(1) Longitudinally, *i. e.*, axial elongation and contraction.

(2) Torsionally, *i. e.*, twist about the axis.

(3) Radially as a whole.

(4) Longitudinal vibrations propagated from any point both ways round the circumference.

(5) Flexurally about diameters after the manner of a disk.

(6) Flexurally in planes at right angles to the axis of the coil.

By calculation using a ring 2·55 cm. internal radius, 1·0 cm. long, and 0·062 cm. thick—these being dimensions approximating to the coils described herein—(1), (2), (3), (4), and (5) yield frequencies outside the important audible range. This leaves the flexural vibrations in the plane of the ring for our consideration, *i. e.*, no. 6.

The modes corresponding to this case can be excited by suspending the ring with a thin silk cord and using a felt hammer or the like as a striker. It is difficult, however, to separate out the various tones. The fundamental—which is characterized by *four* nodes—was obtained in this way, and agreed closely with the calculated value of 440 cycles.

The remaining modes corresponding to this case can be excited by aid of a violin bow. The bow is operated to obtain one note at a time so far as possible. A pure sine-wave valve-oscillator is used to energize a loud speaker, and its frequency adjusted until it gives zero beat with that from the ring. In this way the frequency of the ring mode is determined. It is necessary to mount the ring in some manner and to associate it with an agent whereby the sound

is reinforced. Moreover, the ring was glued to a hollow paper cylinder, 2·55 cm. radius, 0·014 cm. thick, and 1·25 cm. long, the cylinder projecting 0·31 cm. within the ring. The cylinder was glued at its outer end to an annular ring of paper about 0·5 cm. in radial width.

When this structure was suspended by a fine silk thread the ring modes could not be obtained by impact, owing to the large damping effect of the paper. The structure was therefore placed on a large wooden table, held securely by hand (on the annular ring) and bowed, with the results shown in Table VI. When seated on cotton wool only one frequency (520 cycles) could be obtained, and this was quite weak, showing the necessity for the aid of a sounding device.

TABLE VI.

Modes of Copper Ring mounted on Paper Cylinder and stroked with Violin Bow.

Frequency (cycles per second).			Frequency ratio.	
On cotton wool.	On table.	On diaphragm.	Theory (ring free).	Experiment (on table).
520	560	550		
—	1440	1440	2·83	2·6
—	2950	2900	5·42	5·3

The fundamental mode of the ring mounted on the paper cylinder is 560 cycles as against 440 cycles when free. This increase in frequency is evidently due to the stiffening effect of the glue and 0·31 cm. of paper within the ring. The frequency ratios of the various modes are given in column 5, and compare favourably with those obtained by theory * for the ring when quite free. Column 3 shows the frequencies when the ring structure was held in place on a large conical diaphragm. There is little variation compared with the results when the ring was bowed on the table (column 2).

Tests with a formerless coil of 46 S.W.G. enamelled copper wire 2·54 cm. internal radius, baked hard and mounted *viâ* a short thin neck to a diaphragm whose dimensions are stated under Table V., gave the frequencies in Table VII.

* Rayleigh, 'Sound,' i. p. 386.

As a variation on the preceding experiments some data relating to a low-resistance moving coil wound with 32 S.W.G. D.S.C. wire may be of interest. The coil had 55 turns, and its dimensions were similar to those already

TABLE VII.

Formerless Coil of 46 S.W.G. Enamelled Wire, 2·54 cm. internal radius, varnished, baked hard, and mounted on Diaphragm.

Natural frequency of coil (cycles per second).	Frequency ratio.		Remarks.
	Theory.	Experiment.	
690	2·83	2·76	
1900			Most easily obtained with bow.
4000	5·42	5·8	
>4000			Pitch difficult to determine accurately.

TABLE VIII.

Frequencies of Low-resistance Coil attached by Stiff Neck Conical Diaphragm 12·2 cm. radius. Apex angle 90°.

Natural frequency of coil (cycles per second).	Frequency ratio to lowest note.
1300	1
1650	1·27
2800	2·16
3300	2·54

described excepting that it was a trifle longer. The natural frequencies were found by aid of a violin bow, and the results are given in Table VIII.

The ratios in Table VIII., column 2, do not resemble those obtained with a high-resistance coil of 46 enamelled wire as given in Tables VI. and VII. The lowest frequency of 1300 cycles appears, from the previous experiments on the high-resistance coils, to be on the high side, although persistent efforts with the bow failed to elicit any response below 1300 cycles. This may be due to the general

mechanical construction of the coil and its attachment to the diaphragm. The structure was very rigid, and the enhanced rigidity may be responsible for the elevation of the first natural frequency to 1300 cycles. Following the result obtained from the high-resistance coil, the second natural frequency should be approximately $1300 \times 2{\cdot}8 = 3640$ cycles. The nearest frequency to this is 3300 cycles, which is *fourth* on the list. Clearly this coil does not conform with the results obtained from the cylindrical copper ring, since there are two frequencies of 1650 and 2800 interpolated between 1300 cycles and 3300 cycles. The two intermediate frequencies may belong to another form of vibration. There is a longitudinal and a radial repulsive force between adjacent turns which has a maximum value twice per alternation of the current, but oscillations due to this would not be produced by bowing.

Records of the sound emitted whilst bowing the low-resistance coil were obtained. The coil was bowed in the vicinity of a high quality microphone, whose output was amplified and passed on to a recording apparatus. Some of the results are exhibited in figs. 12*a*, 12*b*, 12*c* (Pls. II. & III.). Fig. 12*a* shows the result of a series of sharp flicks with the bow. The effect is to excite not only the fundamental frequency of 1300 cycles, but some of the higher frequencies in addition. In fact, we have here a series of transients, as will be seen by the growth and decay of the oscillation. The record of fig. 12*b* illustrates a frequency of 1600 cycles with a fair degree of purity. The slight variation in amplitude is due to uneven bowing. Fig. 12*c* shows the simultaneous occurrence of two of the higher modes of oscillation, one of which is 2500 cycles. It will be observed that the frequencies found from the records are in close agreement with those obtained by the beat method described herein.

When a coil is used on a diaphragm it is subject to a series of impulses due to transients in the speech or music. For example, when a bass drum is reproduced the coil is suddenly urged backward or forward. Moreover, this impulse will excite all the natural frequencies we have considered above, and the effect will be transmitted to the diaphragm. The extent to which the diaphragm is affected will depend upon the rigidity of the connexion at the top of the coil and also the degree of damping.

9. *Linear Action in Diaphragms.*

Throughout a sound reproducing system it is essential that the action of each component should be linear, *i.e.*, when a

sine wave is applied at the transmitter it should be radiated as a sine wave at the reproducer. In this respect it is of interest to investigate whether the flexural deformation of acoustic diaphragms under static and dynamic forces is linear. By taking a conical diaphragm, loading it at the vertex, and measuring the deflexion either optically or by a screw gauge, the static relationship between load and deflexion can be ascertained. For normal paper diaphragms we have found that the law of linearity holds within a definite limit. For the dynamic case the diaphragm is energized by sinusoidal alternating currents of varying amounts and the axial pressure measured, using the apparatus illustrated in fig. 17. The relationship as plotted is again found to be linear up to a certain limit, which, however, was not exceeded. At the same time it is desirable to examine the wave form of the acoustic output, to ascertain whether it deviates from a sinusoidal shape. This examination was effected by aid of the apparatus of fig. 17, and no distortion was discernible.

In these experiments the diaphragm does not move as a whole except at very low frequencies, so that it is quite different from the static condition. Moreover, the flexure of the diaphragm is divided up into positive and negative zones according to the frequency. The degree of flexure increases with decrease in frequency, so that the stress will be greatest at low frequencies.

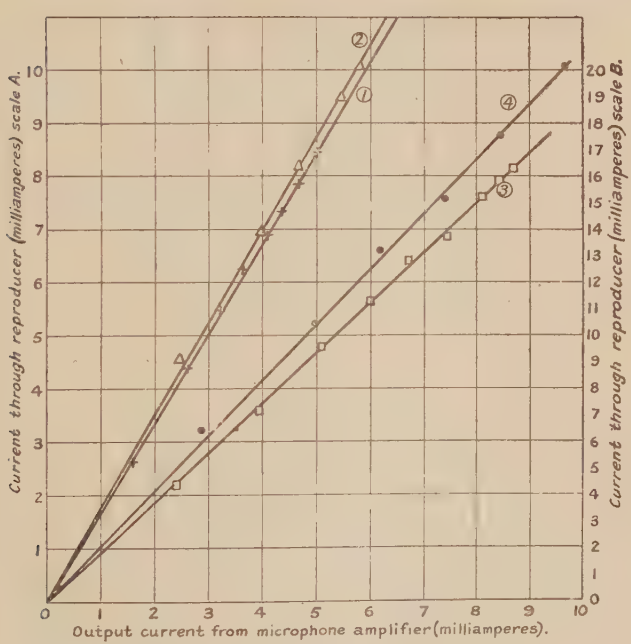
Results of tests of the two different types of drive are given in fig. 13. In one case the force is applied to the apex of the diaphragm (reed), whilst in the other it is applied to a circular ring (coil). The relationship is substantially linear in each case.

If experiments of this nature are conducted in an enclosure where the damping is low there will be standing waves of appreciable magnitude. Provided the physical state of the air remains constant (temperature etc. constant), and the reaction of every part of the room is linear, the standing wave pattern will be unaltered by variation in intensity at any given frequency. Moreover, the position of the acoustic detector at any particular frequency is immaterial. In the experiments described herein there were slight deviations from linearity, doubtless due to atmospheric changes, to temperature rise of the loud speaker coils, or to experimental errors. The experimental arrangement did not permit an examination of the relationship between input and axial pressure when the radiation at low frequencies was large. The relationship was certainly linear for an output of 0·1 watt at 50 cycles, which is adequate for ordinary domestic purposes.

10. *Accession to Inertia of a Vibrating Diaphragm.*

Measurements have been made on the above, using a coil-driven conical diaphragm. The method of procedure is based on an expression which can be deduced from (12) of the former paper*. The reactance of the coil in motion is

Fig. 13.



Curves showing that relation between electrical input and acoustic output of reed- and coil-driven conical diaphragms is linear.

Curve 1.—Coil-driven diaphragm 12 cm. rad., the coil having 2500 turns 48 S.W.G. enamel, tested at 250~(scale A). Force on coil $=9\times10^8$ dynes/abs. unit.

Curve 2.—Coil-driven diaphragm 12 cm. rad., the coil having 2500 turns 48 S.W.G. enamel, tested at 750~(scale A). Force on coil $=9\times10^8$ dynes/abs. unit.

Curve 3.—Coil-driven diaphragm 12 cm. rad., the coil having 2500 turns 48 S.W.G. enamel, tested at 1000~(scale B). Force on coil $=9\times10^8$ dynes/abs. unit.

Curve 4.—Reed-driven diaphragm 21 cm. effective rad., with reinforced apex 4·6 cm. rad., tested at 500~(scale B). Force on reed $=2\times10^8$ dynes per abs. unit.

* Phil. Mag. *loc. cit.*

given by

$$X = \omega L_0 - \frac{1}{\omega C_m}, \quad . \quad . \quad . \quad . \quad . \quad (10)$$

where

$$C_m, \text{ the motional capacity, } \doteqdot \frac{m}{C^2}\text{*}, \quad . \quad . \quad (11)$$

where

$m = M + M_1 =$ natural mass of moving system + accession to inertia,

$C = 2\pi nr H \cos\phi$ (found by measurement) = length of wire on coil × field strength × $\cos\phi$,

$\phi =$ angle between total current and magnetizing current in moving coil †,

$L_0 =$ inductance of coil at rest.

From this we find that the accession to inertia

$$M_1 = \frac{C^2}{\omega(\omega L_0 - X)} - M. \quad . \quad . \quad . \quad . \quad (12)$$

Thus, if measurements of the static inductance and also of the reactance whilst the coil is in motion are made, it is possible to calculate M_1 ‡. There are several precautions to be taken. It is assumed that the cone and coil move as a rigid structure, but this is only true at low frequencies. Consequently the tests should be conducted at low frequencies well below the "break-up" point of the diaphragm, where theoretical and practical conditions are substantially in agreement. A small diaphragm is useful in this case. So long as the distribution of radiation from the cone is reasonably spherical, it is feasible to compare the accession to inertia with that calculated for a rigid disk. We have done this, and find an increase in the mass of the diaphragm of the order of magnitude predicted from Rayleigh's formula §. An unfortunate accident to the apparatus during the course of the work prevented an accurate knowledge of the field strength, and for this reason it is thought advisable not to present numerical data which may be doubtful. It is desirable

* This value of C must not be confused with that for the reed movement.

† See Section 1, where the point has been treated in an alternative way.

‡ Measurements can be made with and without a baffle.

§ Phil. Mag. *loc. cit.*

to measure the static inductance with the field magnet excited, in order to work the iron under the same degree of D.C. polarization *. When so doing it is not easy to fix the coil without moving its position. The simplest case for measuring the accession to inertia is probably where the electromechanical resonance frequency occurs. This can be lowered to come well below the "break-up" of the diaphragm by reducing the strength of the magnetic field.

It has been shown that below a certain frequency the e.m.f. induced in the coil, due to its motion in the magnetic field, causes an effect equivalent to a condenser. Electromechanical resonance occurs when the reactance of the fictitious condenser is equal to that of the coil inductance. For a coil-driven rigid disk the reactance-frequency curve is quite smooth. However, in practice this curve is modified, due to the low velocity of propagation down the diaphragm, thereby giving rise to "modes" at relatively low frequencies. Also due to resonances there may be several frequencies at which the reactance is zero. The e.m. resonance frequency is given by the expression

$$f = \frac{1}{2\pi}\left(\frac{1}{L_0 C_m}\right)^{\frac{1}{2}}, \quad . \quad . \quad . \quad . \quad . \quad (13)$$

where L_0 = static inductance.

$$C_m = \frac{m}{C^2} = \text{motional capacity.}$$

Since C varies directly as the strength of the magnetic field, the e.m. resonance frequency can be raised or lowered accordingly.

11. *Axial Amplitude* $X_{max.}$ *of a Rigid Disk to Radiate a Definite Amount of Power.*

The power radiated in joules per sec. (watts) is given by †

$$W = B\dot{X}^2_{r.m.s.} = B\omega^2 X^2_{r.m.s.},$$

where the axial motion with time is sinusoidal. Thus

$$X_{r.m.s.} = \frac{1}{\omega}\left(\frac{W}{B}\right)^{\frac{1}{2}}. \quad . \quad . \quad . \quad . \quad . \quad (14)$$

* The iron operates on a subsidiary loop.

† Phil. Mag. *loc. cit.* expression (32). In expression (42) of that paper F/m should read $(F/m)^2$.

At low frequencies, when $ka < 0{\cdot}4$, the r.m.s. value of X can be found by substituting †

$$B = \frac{\rho_d \pi a^4}{c}, \quad \text{whence} \quad X_{\text{r.m.s.}} = \frac{9{\cdot}1 \times 10^6 W^{\frac{1}{2}}}{\omega^2 a^2}, \qquad (15)$$

$$X_{\text{max.}} = \frac{1{\cdot}29 \times 10^7 W^{\frac{1}{2}}}{\omega^2 a^2}, \quad . \quad . \quad . \quad . \quad (16)$$

where W is the output in watts at any given value of ω.

TABLE IX.

Showing Axial Movement (maximum value) of a Rigid Disk 10 cm. radius to radiate 1 watt (10^7 ergs./sec.) at various Frequencies.

f. Frequency (cycles per second).	B. Acoustic pressure on disk per unit velocity (dynes).	$X_{\text{max.}}$ Maximum amplitude of disk to radiate 1 watt (cm.).
30	44	$3{\cdot}58$
50	122	$1{\cdot}3$
100	486	$3{\cdot}23 \times 10^{-1}$
200	1946	$8{\cdot}10 \times 10^{-2}$
500	$1{\cdot}05 \times 10^4$	$1{\cdot}39 \times 10^{-2}$
1000	$2{\cdot}7 \times 10^4$	$4{\cdot}33 \times 10^{-3}$
2000	$2{\cdot}7 \times 10^4$ *	$2{\cdot}17 \times 10^{-3}$
4000	$2{\cdot}7 \times 10^4$ *	$1{\cdot}09 \times 10^{-3}$
8000	$2{\cdot}7 \times 10^4$ *	$5{\cdot}44 \times 10^{-4}$

* The oscillation in the value of B is here neglected.

Using formula (14) the axial amplitudes corresponding to the radiation of 1 watt have been calculated, and are given in Table IX.

To avoid distortion, the static and dynamic relationship between axial deflexion and force should be linear over a range adequate to cope with the amplitude requirements at the lowest frequencies. It often occurs in practice, when radiating powerful low notes, that the axial amplitude of a coil-driven diaphragm exceeds the limits over which the surround constraint is elastic. Beyond this point the relationship between force and axial motion soon reaches a saturation value, after which extension ceases. Under this

† Phil. Mag. *loc. cit.* expression (39).

condition a pure note of 50 cycles will be accompanied by strong overtones. Since the sensitivity of the ear increases with rise in frequency (up to about 2000 cycles), the overtones may appear more powerful than the fundamental.

The limits of linearity of the surround can be found for the static case by passing a current through the coil and observing the inward and outward deflexions (according to the direction of the current) of the outer edge of the diaphragm. At the same time the axial extension and contraction of the diaphragm can be ascertained by taking the difference in deflexion between two suitable points, *e. g.*, the outer edge and the edge where the coil is attached.

For the dynamic case the limits of linearity can be found by supplying the coil with a sine wave-current of suitable

Fig. 14.

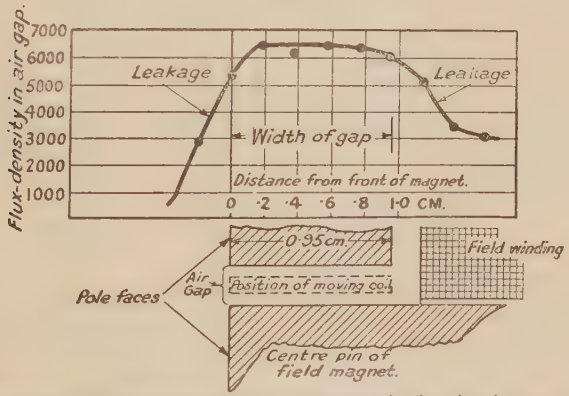


Diagram showing distribution of magnetic flux in air-gap of a circular type magnet used for coil drive.

frequency, and observing, by aid of the apparatus of fig. 17, when the acoustic output departs therefrom.

In practice when the energy radiated exceeds a certain amount, it is clear that the coil will move out of the sensibly uniform part of the field, thereby introducing alien frequencies. A larger diaphragm means a reduction in axial displacement, but there is an upper limit beyond which the output is reduced, due to the break-up frequency being too low (neutralization). Moreover, with the average reproducer it is sometimes essential to reduce the low frequency input to the coil, to avoid mechanical fracture and the production of harmonics (due to limitation of amplitude). The distribution of the magnetic field in the air-gap of a circular magnet with a central pole is portrayed in fig. 14. When

the amplitude of the coil is adequate to bring it beyond the front or rear of the sensibly uniform portion of the field, overtones will be generated. The strength of the field beyond the extremities of the gap indicates considerable leakage. Up to a point this is useful, as it allows (over the sensibly uniform portion) a greater coil amplitude before alien frequencies become prominent. It will be seen that the flux density is sensibly uniform over only 70 per cent. of the width of the air-gap.

From Table IX. we see that the total excursion (twice the maximum amplitude) for frequencies below 150 cycles is very large. For example, at 50 cycles it is 2·6 cm., this being several times the amount obtainable with any reproducer yet constructed. In practice an excursion of this magnitude would not be realized, and the coil and diaphragm would be driven to extreme limits, thereby causing distortion of the type cited above *. For domestic purposes the radiation is usually small and the desired excursions much less than those indicated in Table IX. The latter are closely allied to the operation of reproducers for public entertainment purposes, where the acoustic output is very large.

12. *Amplitude Limit imposed by Mechanical Stress in the Reed.*

Some idea of the axial amplitudes required to radiate low frequencies, *e. g.*, below 200 cycles, will be gained from a cursory glance at the last column of Table IX. in Section 11. These data apply to a disk 10 cm. radius, whereas the disk associated with the reed in Section 1 is 21 cm. radius. In the latter case the amplitudes below 200 cycles will be $\left(\frac{10}{21}\right)^2 = 0{\cdot}23$ those given in Table X. For example, at 50 cycles the maximum amplitude of a disk 21 cm. radius to radiate 0·05 watt is 0·065 cm. or 26 mils (1 mil $=10^{-3}$ inch). From general considerations of elasticity, an amplitude of this magnitude would appear to induce large stresses at the base of the reed where it is rigidly clamped, operating after the manner of a cantilever †. The

* Provided the range of the reproducers were limited to, say, 30 to 200 cycles—by aid of cut-off filter to restrict the input,—larger and heavier diaphragms could be used, thereby necessitating a smaller axial excursion. The range, 200 cycles upwards, would be given by an additional reproducer.

† The data in Tables I. and II. do not refer to a cantilever reed, but to one with a torsional control, *i. e.*, the "Lion" movement.

magnetic force acts—not quite uniformly—over an area at the end of the reed, and causes a bending moment upon it. The centre of magnetic force is slightly nearer the support than the centre line of the pole face.

Taking the necessary data, the stress can be calculated in the usual manner with the following results :—

Width of reed at support	0·65 cm.
Depth of reed at support	0·142 cm.
Distance of c. of m. force from support	1·4 cm.
Stress in material at surface of reed at support	$= \pm 2{\cdot}4 \times 10^{12}$ dynes/cm.²/cm. deflexion. $= \pm 8{\cdot}7 \times 10^{3}$ lbs./in.²/mil. deflexion.
Pull at c of m. force to cause deflexion of 1 mil	$= 8{\cdot}9 \times 10^{5}$ dynes.
at the c. of m. force	= 2 lb.

These results are decidedly illuminating, for it is seen that an amplitude of only 10^{-3} inch is adequate to cause a stress of 8700 lb. per sq. in. in the reed at the point of support, also that the magnetic force required to cause this deflexion is 2 lb. or $8{\cdot}9 \times 10^{5}$ dynes—large figures to be associated with a small instrument.

Considerable interest is attached to a comparison between the power radiated at 50 cycles from a reed-driven disk 21 cm. radius and a coil-driven disk 10 cm. radius, these dimensions corresponding to practical cases. The valve conditions are as follows :—

(1) Internal A.C. resistance of valve = 2000 ohms.

(2) Magnification factor of valve = 2.

(3) R.M.S. voltage change on valve grid = 150.

The power radiated by the reed drive is $0{\cdot}077$ * $\times (0{\cdot}109)^2 = 9{\cdot}2 \times 10^{-4}$ watt, whilst that by the coil drive is 66 † $\times (0{\cdot}07)^2 = 3{\cdot}24 \times 10^{-1}$ watt, or 354 times that from the former. To radiate $3{\cdot}24 \times 10^{-1}$ watt the maximum amplitude of the reed at the point of attachment to the disk (c. of m. force) would have to be 67 mils, or about 1·67 mm. ‡. The stress in the reed corresponding to an amplitude of this amount would be $5{\cdot}8 \times 10^{5}$ lb. per sq. in., which is beyond the elastic limit of the material. The coils would burn out, due to excessive heating caused by a current of 2 amperes r.m.s., yielding an ohmic loss of 3 kilowatts. These data prove beyond doubt that it is impossible to radiate large power at the lower frequencies by a simple flat cantilever

* See Table II. column 2 above.

† See Table II. column 5, p. 1028, Phil. Mag. *loc. cit.*

‡ The current-deflexion relationship would not be linear for this amplitude.

reed-driven apparatus designed to cover a wide frequency band, whilst (*a*) obtaining a good upper register, (*b*) preserving a linear current-deflexion relationship, (*c*) keeping the maximum stress in the reed below the elastic limit of the material, (*d*) avoiding undue heating of the operating coils.

The power could be radiated if the reed were more flexible (lower natural frequency) and the response restricted to the lower and middle frequencies. The upper frequencies would, of course—as in the case stated previously—be supplied

Table X.

Radial Velocity of Sound in a Circular Aluminium Disk.

Radius of disk=10 cm.; thickness of disk=0·165 cm.

f. Frequency (cycles per second).	Radius of node (cm.).	Radial distance between consecutive nodes. $\lambda/2$ (cm.).		Mean radial velocity between consecutive nodes. v_r (cm. sec.$^{-1}$).	
		Experiment.	Calculation.	Experiment.	Calculation.
590	7·5				
1610	3·75				
		4·5	5·0	$1{\cdot}45\times10$	$1{\cdot}61\times10^4$
	8·25				
3320	1·97				
	5·5	3·53	3·52	$2{\cdot}34\times10^4$	$2{\cdot}33\times10^4$
	8·75	3·25	3·52	$2{\cdot}16\times10^4$	$2{\cdot}33\times10^4$

by another reproducer. In fact, with both reed- and coil-driven diaphragms it seems essential to use more than one unit to increase the range of response beyond that given by modern commercial apparatus.

The reed stress due to the permanent magnet must not be overlooked, and in designing a movement it must be borne in mind that the maximum force is the sum of the effects due to the current in the coils and to the permanent magnet. In this respect, therefore, it is imperative to limit the forces due to both of these attractions.

13. *Radial Velocity of Sound in a Flexible Disk.*

In Section 3 data were given relating to the modes of a flexible disk driven by a reed whose natural frequency was about 2600 cycles per second. It was shown that an increase

in the frequency range was accompanied by an increase in the number of circular nodes on the disk. The radial distance between consecutive nodal circles can be regarded as half the wave-length of sound travelling radially in the disk from its centre. Owing to the stiffness being greatest at the centre the radial velocity will be greatest there also, and will decrease towards the edge. Moreover, the velocity ascertained from nodal measurements can be considered to be an average or mean value for the annulus in question.

TABLE XI.

Radial Velocity of Sound in a Steel Disk.

Radius of disk=9·4 cm.; thickness of disk=0·058 cm.

f. Frequency (cycles per second).	Radius of node (cm.).	Radial distance between consecutive nodes. $\lambda/2$ (cm.).		Mean radial velocity between consecutive nodes. v_r (cm. sec.$^{-1}$).	
		Experiment.	Calculation.	Experiment.	Calculation.
380	6·8				
970	4·16				
	7·56	3·4	3·8	$6·6 \times 10^3$	$7·4 \times 10^3$
1600	2·33				
	5·1	2·77	2·96	$8·9 \times 10^3$	$9·5 \times 10^3$
	7·8	2·7	2·96	$8·6 \times 10^3$	$9·5 \times 10^3$
2540	1·08				
	3·53	2·45	2·34	$1·25 \times 10^4$	$1·19 \times 10^4$
	5·92	2·39	2·34	$1·21 \times 10^4$	$1·19 \times 10^4$
	8·0	2·08	2·34	$1·06 \times 10^4$	$1·19 \times 10^4$

Taking the usual expression $v_r = \lambda f$, where v_r = radial velocity, λ = wave-length, and f = frequency, the data in Tables X. and XI. have been deduced from measurements on an aluminium and on a steel disk.

The calculated values of nodal distances were obtained by aid of the formula* $\lambda = \frac{2\pi}{h}$. Since h varies as $f^{\frac{1}{2}}$ the velocity $v_r = \frac{2\pi f}{h}$ should increase as $f^{\frac{1}{2}}$, and at the higher frequencies this is found to be substantially true for the portions of the

* See Warren, Phil. Mag. ix. p. 900 (May 1930).

disks remote from the centre. The velocity decreases with increase in radius.

The agreement between theory and practice is closer for the steel than for the aluminium disk. The discrepancies in the latter case are doubtless due to some variation from theoretical conditions. This disk was much more rigid than the steel one, and a greater input was required to yield an amplitude adequate to cause the movement of sand to the nodes. Also the values of E and σ were not measured, but extracted from a table of physical constants. On the whole the agreement between theory and practice is adequately close to show that the radial velocity of propagation in any given disk can be calculated with reasonable accuracy, provided the practical conditions do not violate those contemplated in obtaining the analytical expression.

In each of the cases cited above the radial velocity is less than that of sound in air ($3{\cdot}4 \times 10^4$ cm. sec.$^{-1}$ at 20° C.). Thus the radiation from the centre of a disk reaches the edge before the radial sound waves travelling in the material of the disk. Consequently the radiation from a flexible disk is of a more complicated character than that associated with a rigid disk. The axial nodes and the polar diagrams of air-pressure distribution will differ in the two cases.

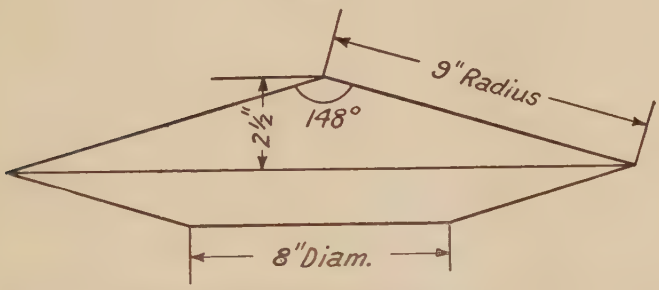
14. *Radial Velocity of Sound in a Conical Diaphragm.*

In order to obtain approximate data relating to the velocity of propagation in a conical paper diaphragm measurements were made during the course of some experiments on dust figures. The dimensions of the diaphragm in question are given in fig. 15. A series of rings were traced by lycopodium powder between a circle of about 10 cm. radius and the periphery. In Table XII. there is a series of values of the nodal distances at various frequencies. The distance between the rings at any given frequency was not quite constant, but a mean value has been taken. The radial velocity of propagation along the outer part of the diaphragm was calculated as in Section 13. Near the apex it was doubtless somewhat greater than the figures given in Table XII. The velocity of sound in air at 20° centigrade is about $3{\cdot}4 \times 10^4$ cm. per second. Thus the velocity of flexural waves down the outer part of the diaphragm at a frequency of 1600 cycles is about 1/5 the velocity of sound in air.

From the last column of Table XII. we see that the

velocity of propagation down the diaphragm increases with increase in frequency approximately in accordance with the relationship $v_r \propto f^{\frac{1}{2}}$, as in the case of flexible disks. It should be observed, however, that this relationship has not been verified for points near the apex. This increase in velocity is akin to that in a telephone cable, where the higher

Fig. 15.



Dimensions of diaphragm used for determination of radial velocity. Thickness of (soft) paper 0·02 inch.

TABLE XII.

Showing Wave-length and Radial Velocity of Propagation in Paper Diaphragm 0·05 cm. thick at various Frequencies.

Velocity of sound in air at 20° centigrade $= 3{\cdot}4 \times 10^4$ cm. per second.

Frequency (cycles per second).	Mean distance between nodal circles (cm.).	Wave-length on diaphragm (cm.).	Velocity in diaphragm (cm. per sec.).
1600 *	2·2	4·4	$7{\cdot}0 \times 10^3$
2000	1·9	3·8	$7{\cdot}6 \times 10^3$
2900	1·5	3·0	$8{\cdot}7 \times 10^3$

* Accompanied at this frequency by radial nodes.

frequencies travel more quickly but are more heavily attenuated than the lower frequencies.

15. *Reproduction of Transients.*

Analyses associated with the reproduction of sound usually contemplate the steady state. As speech and music consist

mainly of transients, this side of the subject deserves considerably more attention than it has been afforded hitherto. The conditions under which a transient can be perfectly reproduced have been indicated previously in this Journal *. Such conditions are never enjoyed in practice, and it is necessary therefore to deal with cases which, if not absolutely practical, are closely allied to practical conditions. A transient, whatever its shape, can be expressed mathematically by a Fourier integral. It consists of an infinite spectrum of frequencies ranging from zero. Not only are the amplitudes of the component frequencies important, but also their phases. If we consider an acoustic impulse which rises very rapidly to a maximum value, it is clear that the higher frequency components of the spectrum are of paramount importance. If, due to inductance in the electrical portion of a reproducing system, the higher frequencies are delayed, the initial sharpness of the impulse is reduced. This is due to the alteration in phase and reduction in amplitude of the component oscillations constituting the impulse.

By aid of the data in the earlier paper, and those relating to the focussing effect of a rigid circular diaphragm at the higher frequencies, it is possible to calculate the profile of an impulse reproduced by a coil-driven disk. We shall postulate a disk 10 cm. radius driven under conditions stipulated in the earlier paper. An acoustic detector is situated on the axis of the disk, and is assumed to leave the sound field undisturbed. An electrical wave-form is applied to the grid of a thermionic valve and thence to the coil. The detector is situated at such a distance from the disk that the resultant radiation from different parts thereof arrives substantially in the same phase at all frequencies. It is desired to find the acoustic wave-form of the transient at the detector. The acoustic pressure at the detector is directly proportional—for set conditions—to the acceleration of the disk.

A simple wave-form consists of a series of regular interruptions (Morse dots) whose profile is strictly rectangular. Although this is not a transient in the rigorous sense, it serves to illustrate—particularly in practice—the problem under consideration.

The wave-form is given by the Fourier series :—

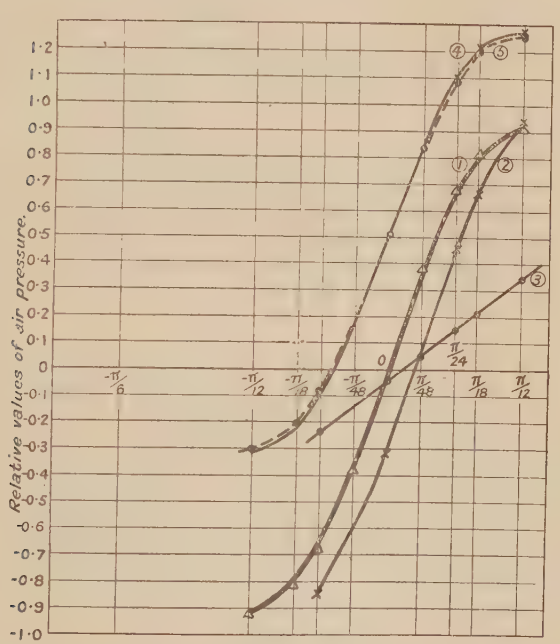
$$y = \mathrm{P}\{\sin \omega t + \tfrac{1}{3}\sin 3\omega t + \tfrac{1}{5}\sin 5\omega t + \text{etc.}\} \quad . \quad . \quad . \quad (17)$$

To save labour in computation it is proposed to use only six terms of this series. Provided ω is adequately high, the

* Phil. Mag. *loc. cit.* p. 1036.

profile of the dots is approximately rectangular, although, of course, this is largely an optical effect depending upon the scale of the time base. However, with a large value of ω and a reasonable amplitude, the wave-front is steep so far as aural appreciation is concerned, and this is preeminently the aspect of the analysis in which we are immediately interested.

Fig. 16.



Curves showing air pressure on axis and at 45° to axis, corresponding to a voltage wave-form indicated by expression (17), as reproduced by a coil-driven rigid disk.

Curve 1. Voltage wave-form applied to grid of power valve.
Curve 2. Air pressure on axis with dot frequency of 500 per second.
Curve 3. Air pressure at 45° to axis with dot frequency of 500 per second.
Curve 4. Air pressure on axis with dot frequency of 50 per second.
Curve 5. Air pressure at 45° to axis with dot frequency of 50 per second.

16. *Graphical Illustration of Wave-form cited in Section* 15.

Curves computed by aid of (17) showing the axial air pressure many diameters distant from the disk under different conditions are given in fig. 16. Curve 1 shows the profile

of the dot as applied to the grid of the valve, and indicates what the acoustic reproduction should be if it were perfect. Curves 4 and 2 illustrate the axial pressure for fundamental dot frequencies of 50 and 500 cycles per second respectively. The former is shifted backwards in phase and upwards relative to the zero line. This is due to the motional capacity advancing the phase of the lower frequency components and altering their relations. The 500 cycles curve 2 lags on curve 1, owing to inductance. The "lag" is not so prominent here as the "lead" of curve 4, since the fundamental frequency in the latter case is further away from the electromechanical resonance frequency, *i. e.*, about 350 cycles where the reactance is zero.

There is little difference between the *profiles* of the curves of axial pressure at 50 and 500 ~. Moreover, in listening to a rhythmic series of dots it is probable that there would be little distortion apparent to the ear. The effect, however, in a solitary transient of the phase shift evinced in curve 4 might be perceptible, but this could only be settled experimentally.

The air pressure at an angle of 45 degrees to the axis is illustrated in curves 3 and 5. It is clear from curve 3 that with a dot frequency of 500 cycles we are faced with evanescence of the higher frequency components due to interference, whereas at 50 cycles (curve 5) the higher frequencies are substantially equal to those on the axis.

Obviously unless the ear were driven to saturation by abnormal acoustic pressure due to large amplitude the difference between curves 4 and 3 or curves 1 and 3 would be quite easily perceived.

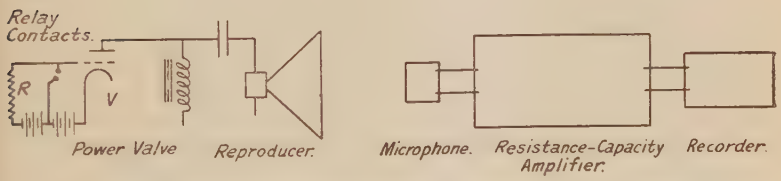
17. *Comparison between Theoretical and Actual Transients.*

It is of considerable scientific interest and of great practical importance to compare the profiles of fig. 16 with those obtained from actual reproducers. Several differences in the conditions between the two cases are to be noted. A fairly close approximation to a rectangular wave-form can be applied to the grid of the valve, but the detector—in this case a microphone—has an effect upon the sound field, particularly the strength of the higher frequency components, which will in general be increased, due to diffraction. However, we are not in quest of accurate quantitative data, since the diversity between hypothesis and reality is too grotesque to demand rigorous accuracy. The general arrangement of the apparatus for determining the acoustic wave-form

corresponding to Morse dots is shown in fig. 17. A special microphone, having no appreciable resonance within the audible register, is situated on the axis of the loud speaker in a draped enclosure. The contacts of a relay which operates about 15 times per second are used to short-circuit a resistance R which alters the grid bias on the valve. This causes a variation in the anode current of the valve. The reproducer is situated in the anode circuit of V, and it responds to the variations in anode current. Owing to the reactance of the loud speaker the profile of the current wave-form in its windings is slightly rounded. The output from the reproducer is registered by the microphone, amplified by a resistance capacity amplifier, and revealed by a suitable recording device.

Records have been taken for a number of reproducers, all of which exhibit natural oscillations whose damping depends upon the type of reproducer. Fig. 18 (Pl. I.) is the record

Fig. 17.



Apparatus for determining the acoustic output from a reproducer.

for the reed-driven aluminium disk cited in section 3. Instead of the rectangular wave-form being reproduced we have a complex damped oscillation. Viewed broadly the oscillation consists of one or more high-frequency components superposed on a main low-frequency oscillation of very small damping. The frequency of the main oscillation is 560 cycles per second, and this corresponds closely to the second mode of the combination of reed and disk, viz., 580 cycles. The third mode is also in evidence on the record. The first mode occurs at 144 cycles, but there is no trace of this, since it is a "fixed" centre condition and the impulse was applied at the centre, thereby causing a central displacement. In any case the energy of the fixed centre oscillations is small.

The record for a reed-driven paper diaphragm (the "Lion" L.S. of section 2) shown in fig. 19 (Pl. I.) indicates a severer degree of damping than in the case of the aluminium disk. This is one reason why the paper diaphragm is a vastly superior reproducer of speech and music. An analysis of the

main frequency components is of utility from a practical view-point. The frequencies of the two main oscillations are 2000 and 230 cycles respectively. The higher-frequency oscillation corresponds to that of the reed, clamping nuts and screwed stem vibrating as a whole. Thus, at this frequency the apparent mass of the diaphragm is zero *. Reference to fig. 19 (Pl. I.) shows that the most prominent oscillation occurs at a frequency of 230 cycles. Although the variation in effective inductance and resistance is much less marked at 230 cycles than it is at 116 cycles (see fig. 6), it is obviously a second main mode of the diaphragm-reed combination. This case is similar to that of the aluminium disk quoted above. The 116 cycle mode is doubtless a "fixed centre" type, and thus a central impulse does not excite it appreciably, which is a fortunate practical circumstance. A point of importance is that although the conical diaphragm has a multitude of modes only one of these is prominent, *i. e.*, 230 cycles. There are other modes less clearly defined. For example, a mode is indicated between 50 and 60 cycles, and this can be seen on the record.

A record taken with a transformer coupled coil-driven conical diaphragm is illustrated in fig. 20 (Pl. III.). Again, there are two main oscillations of high and low frequency respectively. The heavily damped low frequency is due to the rubber surround *per se* (section 7), whilst the high frequency is due to the elasticity of the diaphragm. In transients of this nature the influence of the low velocity of propagation in the diaphragm plays its part in moulding the profile of the sound record. The 2000 $\sim$ radiation from the apex will reach the microphone several cycles before that from the periphery.

For completeness, records were obtained of two important cases relating to a reproducer of the horn variety. The driving agent was of the moving coil type, identical with that used for talking pictures, whilst the horn was an exponential one with straight axis 5 feet long. Both with and without the horn the wave-forms were similar in nature to those depicted already. Two frequencies of about 1500 and 300 cycles were present without the horn. The lower of the two is the natural frequency of the diaphragm loaded by the coil, whilst the higher is doubtless another diaphragm mode. The damping effect of the horn was clearly shown, and the 300 cycle mode was almost extinguished. The horn itself †,

* It might be negative (see fig. 4).
† Not the air column.

when tapped, had a main natural frequency about 200 cycles, and this was prominent on the record.

All the preceding tests were conducted in a draped enclosure, to avoid standing waves etc., so that merely the direct output due to the impulse was recorded. It is hardly necessary to remark that the superposition of these natural oscillations upon speech and music results in acoustic coloration.

18. *Causes of Distortion in a Moving Coil System.*

Before recounting the various causes of distortion it will be advisable to attempt to define what is meant by the term. Broadly speaking there is acoustic distortion when the air pressure due to the reproducer is not a scale replica of that in the broadcasting studio. This, however, requires to be more detailed to bring it within our immediate scope, since in general there will be distortion due to electrical causes before the voltage is applied to the power valve. There will also be distortion due to reflexion from the walls, furniture, etc. of the room where the reproducer operates. This will be accentuated if the amplitude of any of the room components is large enough to introduce non-linear action. When the reproducer operates in free air, if we imagine the voltage applied to the power valve (assumed to have infinite input impedance) to be directly proportional to the air pressure in the studio, and the various phases in the electrical and acoustic waves to correspond, then anything which modifies either the relative amplitudes or the phases or introduces alien frequencies can be regarded as a source of distortion. The causes of distortion can be classified under three heads :—(*a*) Electrical ; (*b*) Acoustical ; (*c*) Mechanical.

Electrical.

1. Variation in the impedance of the moving coil throughout the frequency range *.

2. Frequency multiplication or generation of harmonics, due to mutual inductance between the coil and the magnet †. The alternating current in the coil magnetizes the field magnet, and the coil thereby moves in a steady field with an A.C. field superposed.

* Apart from variations due to resonances, there is the effect of the motion of the coil in the magnetic field and of variations in inductance and effective resistance due to iron loss.

† See Table XIV.

3. Generation of harmonics, due to hysteresis in iron circuit of magnet.

4. Large amplitudes at low frequencies cause the coil to move out of the region where the radial magnetic field is uniform. This also causes a variation in the inductance and effective resistance of the coil. In each case alien frequencies are created.

Acoustical.

5. Non-uniform sound-pressure distribution and reduced energy over the hemispherical surface on one side of the baffle, *i.e.*, focussing due to interference.

6. Accession to inertia at low frequencies, causing a reduction in output.

7. Baffles of inadequate dimensions, thereby causing a reduction in the intensity of the lower frequencies.

Mechanical.

8. Resonance at some frequency below 100 cycles, due to the diaphragm moving as a whole on its elastic support.

9. Symmetrical "modes" (circular) at various frequencies, due to the elasticity of the diaphragm.

10. Radial diaphragm "modes" throughout the frequency range *.

11. Resonance of the annular diaphragm support.

12. Radial "modes" of the moving coil *per se*.

13. Resonance of the baffle used to isolate the two sides of the diaphragm.

14. Reduced radiation at higher frequencies, due to diaphragm transmission loss.

15. Non-linear action (not necessarily asymmetric) of the annular diaphragm support with large amplitudes which may occur at low frequencies. The harmonics will fall where the ear is relatively sensitive, and may appear louder than the fundamental.

As we have shown above, resonances in the upper musical register are essential in the reproducers discussed herein. In a rigorous sense they are a source of distortion, although the

* If the outer edge is serrated, bent over, and glued to the outer supporting annulus, resonance sometimes occurs accompanied by buzzing. It should be understood that in general the diaphragm executes flexural, longitudinal, and torsional oscillations.

acoustic effect transcends that which would be obtained with a rigid disk (excepting on the axis of the latter).

16. Non-linear action of a diaphragm, if the stress is excessive, due to a large axial motion and the radiation of considerable power.

Appendix.

Table XIII.

ka.	G.	$K_1(z)$.
0·25	0·031	0·025
0·5	0·12	0·199
0·75	0·25	0·66
1·0	0·423	1·29
1·5	0·774	3·06
2·0	1·033	4·28
2·5	1·131	4·039
3·0	1·092	2·869
3·5	1·001	2·46
4·0	0·942	3·20
4·5	0·946	7·10
5·0	0·999	9·08
5·5	1·032	8·86
6·0	1·037	6·90
6·5	1·011	5·63
7·0	0·981	6·62
7·5	0·973	9·94
8·0	0·989	13·06
8·5	1·011	13·86
9·0	1·021	12·0
9·5	1·011	9·18
10·0	0·993	9·26

$ka = \dfrac{\omega a}{c}$ where $\omega = 2\pi \times$ frequency, $a =$ radius of disk, and $c =$ velocity of sound in air.

$$G = \left[1 - \frac{J_1(2ka)}{ka}\right]. \quad . \quad . \quad . \quad . \quad . \quad (18)$$

For small values of z the series used is

$$K_1(z) = K_1(2ka) = \frac{2}{\pi}\left[\frac{z^3}{1^2.3} - \frac{z^5}{1^2.3^2.5} + \frac{z^7}{1^2.3^2.5^2.7} - \text{etc.}\right]. \quad . \quad . \quad . \quad . \quad (19)$$

For large values of z, by substituting multiples of $\frac{\pi}{4}$ for z, the series used is

$$K_1(z) = \frac{2}{\pi}\left\{z+z^{-1}-3z^{-3}+1^2.3^2.5z^{-5}-1^2.3^2.5^2.7z^{-7}+\ldots\right\}$$
$$-\left(\frac{2z}{\pi}\right)^{\frac{1}{2}}\left[\cos\left(z-\frac{\pi}{4}\right)\left\{1-\frac{(1^2-4)(3^2-4)}{1.2.(8z)^2}\right.\right.$$
$$\left.+\frac{(1^2-4)(3^2-4)(5^2-4)(7^2-4)}{1.2.3.4.(8z)^4}-\cdots\right\}$$
$$+\sin\left(z-\frac{\pi}{4}\right)\left\{\frac{1^2-4}{1.8z}-\frac{(1^2-4)(3^2-4)(5^2-4)}{1.2.3.(8z)^3}+\cdots\right\}\Bigg] \quad . \quad . \quad . \quad (20)$$

TABLE XIV.

Illustrating the mutual inductance between the fixed (magnetizing) and moving coils of a reproducer, both coils being clamped. In practice, since one coil is in motion, the conditions would be modified.

The moving coil had 1000 turns 5 cm. diam., 1 cm. long.

Frequency (cycles per second).	Effective inductance of moving coil (henry).	Condition.
50	0·29	Magnetizing coil open (no excitation).
50	0·25	Magnetizing coil closed (no excitation).
50	0·27	0·16 amp. in magnetizing coil of 15,000 turns.

May 1930.

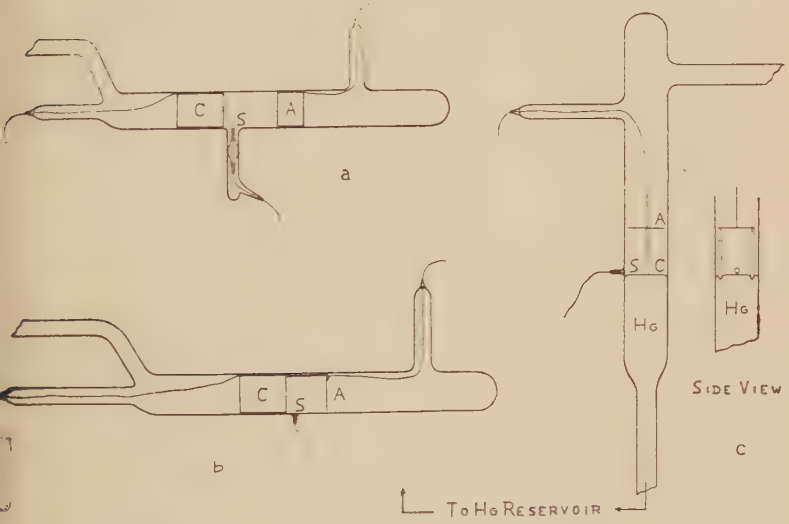
II. *The Potential of the Walls in the Cathode Dark Space.* *By* JEAN W. BECK, *M.Sc., and* K. G. EMELÉUS, *M.A., Ph.D., Department of Physics, The Queen's University of Belfast* *.

A METHOD for finding potentials in the cathode dark space has recently been developed by W. L. Brown and E. E. Thomson [1] which affords a test of Morse's theory of the dark space and leads to a determination of

* Communicated by the Authors.

Townsend's ionization coefficients for electrons. Brown and Thomson's values for these differ somewhat from Townsend's, and it was thought that this might be due in part to the effect of wall charges in the discharge-tubes used by them. For this and other reasons it seemed of interest to undertake the investigation of the wall potential described in the present paper. It has been necessary, in view of the enormous range of discharge conditions possible, to limit the investigation to conditions generally similar to those employed by Brown and Thomson and others working in this laboratory. The results obtained, although

Fig. 1.



Types of discharge-tubes.

quite definite in their main features, have proved to be irregular in detail. We have not been able to find with certainty the causes of these irregularities, but suspect that they arise from the effects of impurities in the tubes ; it is questionable if a more extended investigation is worth while at present.

1. *Experimental Arrangements.*

Medium-sized tubes of soda-glass similar to those used by Brown and Thomson have been employed. The first type is shown in fig. 1 *a*. The main part of the tube had a diameter of 2·6 cm. The collector (S) was a piece of

0·3 mm. copper-clad wire, sealed into a glass sheath. The cathode (C) was a hollow iron cylinder 3·5 cm. long, 2·5 cm. wide, closed at the end exposed to the discharge by a disk of nickel. The anode (A) was a hollow iron cylinder 2 cm. long, open at both ends. Anode and cathode were degassed in a vacuum furnace before assembly of a tube. The two electrodes were moved separately by a magnet, and the distance between them was not accurately constant. A second type, with which most of the measurements were made, is shown in fig. 1 *b*. It had the same cathode as the first, but for anode a disk of nickel connected to the cathode by a glass rod. The collector (S), of 0·3 mm. copper-clad wire, was sealed in by heating a very small spot of the glass wall, pulling it out with a fine point of glass, inserting S in the orifice so formed, and then sinking the glass down as nearly as possible to its original cylindrical form internally. The inner end of S was ground down level with the glass. A third tube is shown in fig. 1 *c*. Here a mercury cathode of the barometer type was used, to find if the glass walls behind the cathode had any effect in 1 *a* and 1 *b*. The main part of the tube had a diameter of 3·2 cm. The anode (A) was a disk of nickel floating on a light glass frame at a distance of about 3 cm. from the cathode. S was inserted as in 1 *b*.

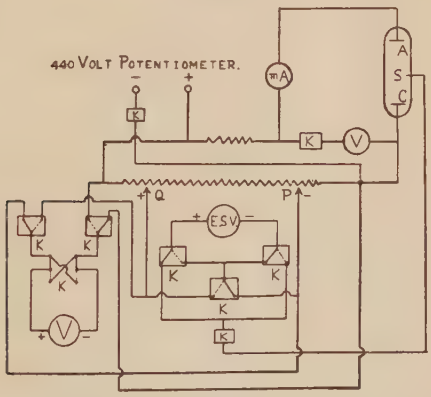
The tubes were exhausted by a single-stage Gaede diffusion pump backed by an oil pump. The discharge tube was connected directly to one liquid air trap, and two further traps were put in the gauge and pump leads respectively. Pressures were recorded by a McLeod gauge. The gases studied were argon, containing 0·1 per cent. of nitrogen ; nitrogen, from sodium azide ; and oxygen, from potassium permanganate. Before admitting each specimen of gas the tubes were baked out at 370° for six hours, unless otherwise stated.

The circuit ultimately used is shown in fig. 2. It allowed the potential taken up by the collector to be measured relative to either anode or cathode, the potential being shared between an electrostatic voltmeter and a Weston moving coil voltmeter. The discharge was run from a potentiometer with a maximum of 440 volts. The procedure consisted in placing the cathode at various distances from the collector, measured with a travelling microscope, and for each position noting the tube voltage, the tube current, the potential assumed by the collector,

the part of the discharge in which the collector was situated, and the approximate width of the cathode dark space. Potential readings for the collector were taken relative to both anode and cathode, a check upon the insulation being afforded by the sum of the two readings being equal to the tube voltage, generally to within 10 volts.

It is admittedly an assumption that the potential taken up by the collector is the same as that of the inner surface of the piece of glass which it replaces. Both glass and metal surfaces will, however, rapidly become covered with a thin film of sputtered material from the cathode when the discharge starts. Also it was found that touching the

Fig. 2.



Electrical circuit.—The potential of S is referred to that of C by connexions indicated by full lines and to that of A by connexions indicated by broken lines. In the first instance P is used, in the second Q.

K, key; E.S.V., electrostatic voltmeter; V, moving coil voltmeter; mA, milliameter.

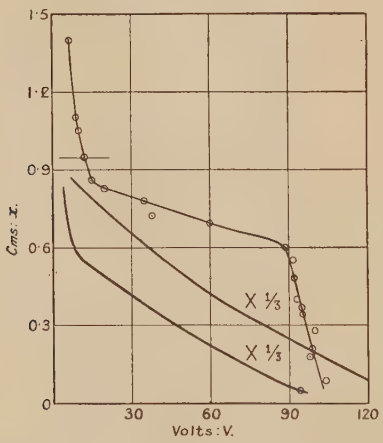
outside wall had no effect on the discharge or on the voltmeter readings, whilst a piece of tinfoil wrapped round the tube in various positions, and held at various potentials, was likewise without effect.

2. *Results.*

The results obtained are illustrated by the curves of figs. 3–8 (taken from 63 in all), in which the collector potential (V) negative to the anode is plotted against distance from the cathode (x) for various conditions of

discharge. The position of the edge of the cathode dark space, judged visually, is shown by a dash on each curve. When the points for a curve are indicated by two different symbols, one refers to readings taken relative to the anode and the other to readings taken relative to the cathode. The curves exhibit less regularity than could be desired, but this can be readily accounted for by the difficulty of obtaining definite conditions of the inner surface of a discharge-tube. In all cases an anode glow was absent, the anode being in the Faraday dark space, so that the potentials are closely the same as if they had been referred to the potential of the middle of the negative glow as zero. On comparing with Brown and Thomson's curves it will

Fig. 3.



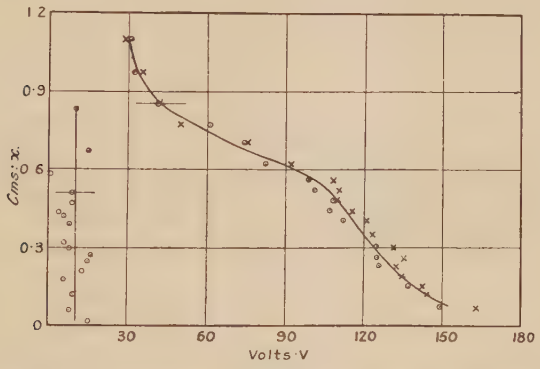
Right-hand curve.—Potential of walls in nitrogen. Tube 1 *a*. Tube voltage 420; tube current 1·78 ma.; pressure 0·3 mm. Hg.
Left-hand curve.—Difference in potential between discharge and walls, potentials to one-third scale.
Middle curve.—Potential of the discharge calculated from Brown and Thomson's run 18, also to one-third scale.

be seen that in the main part of the cathode dark space the walls are strongly positive to the discharge. In the negative glow they are known to be negative, therefore in some narrow region near the negative glow the walls and the discharge are at the same potential. In fig. 3 the potential of the discharge has been calculated from the constants of Brown and Thomson's run 18 (middle curve), and the difference in potential between the discharge and walls also

plotted (lowest curve). The potential difference amounted to about 320 volts immediately against the cathode.

Baking out affected the readings to an important extent only after the tube had been standing for a week or so without liquid air on the traps. When a set of readings was then taken without baking and without liquid air on the traps a rather irregular curve was obtained within a range as low as 20 volts negative to the anode ; in fig. 4 the left-hand line is an example of one of these. After liquid air had been replaced and the tube baked the usual type of curve was obtained (fig. 4, right-hand curve). The tubes were never more than slightly clouded by sputtering, and this was found not to affect the results appreciably.

Fig. 4.



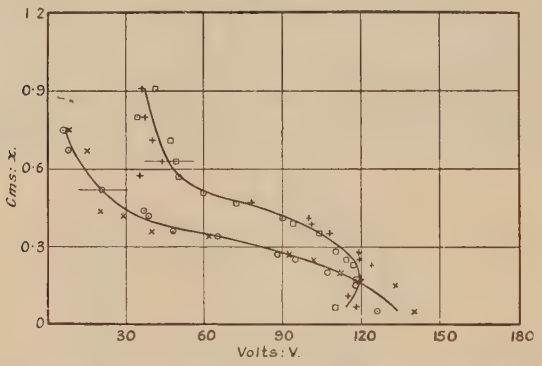
Potential of walls in nitrogen.—Tube 1 *b*, left-hand line, before baking. Tube voltage 431 ; current 1·14 ma. ; pressure 0·44 mm.
Right-hand curve, after baking.—Tube voltage 422 ; current 0·27 ma. ; pressure 0·425 mm. Compare comment on currents in fig. 5.

Tubes of types 1 *a* and 1 *b* gave similar results, but there was a marked difference with the tube 1 *c*. The wall potential in this case was nearer to that of the cathode than with 1 *a* and 1 *b* in the negative part of the dark space, and it appeared as if the wall potential would be that of the cathode at the surface of the latter. With the other tubes, in which there was much free glass behind the cathode, and presumably poor contact between the cathode and the wall, the curves tended to curl round towards the axis of voltage as the cathode was approached. These differences are brought out by a comparison of figs. 5 and 6, which refer to discharges through nitrogen with tubes 1 *b* and 1 *c* respectively, and the close approach

to the cathode potential is well shown on the curve for the lowest tube voltage in fig. 8 for a discharge through argon.

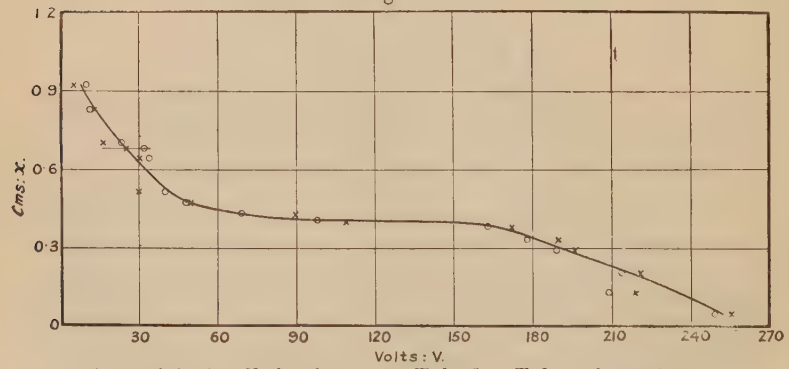
We can trace no definite dependence of the form of the curves on the nature and pressure of the gas and on the tube voltage, a result which may be compared with the similarity of the potential curves for the discharge itself under

Fig. 5.



Potential of walls in nitrogen.—Tube 1 *b*. Tube voltage 425; pressure 0·52 mm. For the left-hand curve the current was 0·775 ma., for the right-hand 0·33 ma. The cause of the difference in the currents was not established.

Fig. 6.

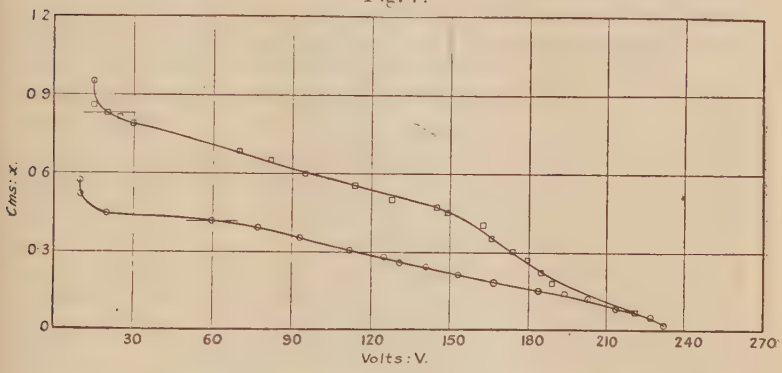


Potential of walls in nitrogen.—Tube 1 *c*. Tube voltage 429; current 0·54 ma.; pressure, 0·41 mm.

varying conditions [1]. With oxygen, however (see fig. 7), there is usually a very rapid initial change in potential in going from the negative glow to the dark space through the boundary region investigated by N. M. Carmichael [2]. Three sets of readings taken with argon at a definite pressure

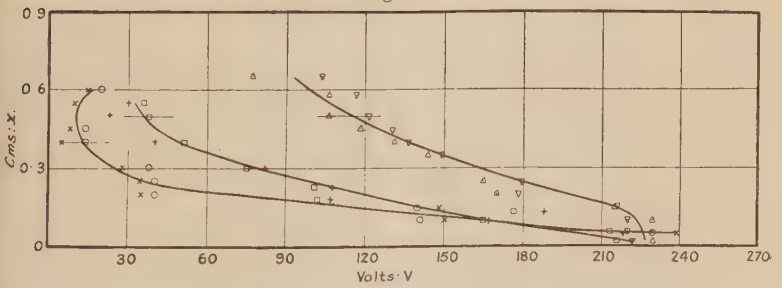
(0·65 mm.) and various tube voltages with tube 1 *c* are shown in fig. 8. The right-hand curve, for the lowest potential, was evidently not followed to the negative glow, but the edge of this much misjudged, on account of its diffuseness.

Fig. 7.



Potential of walls in oxygen.—Tube 1 *b*. Left-hand curve. Tube voltage, 413; current 0·83 ma.; pressure 0·44 mm.
Right-hand curve.—Tube voltage 419; current 0·24 ma.; pressure 0·27 mm.

Fig. 8.



Potential of walls in argon.—Tube 1 *c*. Pressure 0·65 mm.
Bottom curve.—Tube voltage 352; current 2·45 ma.
Middle curve.—Tube voltage 301; current 0·78 ma.
Top curve.—Tube voltage 249; current 0·20 ma.

3. *Discussion of Results.*

It appears impossible at present to give a theory of the dark space in which allowance is made for the effect of the walls. A full account would have to extend analysis similar to that developed by Morse[3], with introduction of the boundary condition that the total current to each element of the wall from the gas is zero, an assumption in

itself doubtful in view of the possible transport of current along surface films. In an approximate treatment Morse's solution could be assumed for the axial part of the tube, and the potential of the walls calculated which would make the net lateral current of positive ions and electrons zero. In either case it would be essential to know the laws governing the scattering of ions and electrons in the dark space, and the direction and speeds of particles after an ionizing collision had occurred, since there is evidence that the ions and electrons follow the lines of force [4] fairly closely; the lateral current would thus be determined largely by the lateral component of the initial directions of motion after collisions. Whilst there is some information available in this connexion for electrons, very little is known about positive ions [5].

The variation in the wall potential can nevertheless be correlated qualitatively with the fraction of the current carried across the dark space by positive ions and electrons. It is sufficient to make use of the theory of the normal cathode fall developed by Compton and Morse [6], since the cathode falls of the present investigation are not much above their normal values. It is believed from this that in argon and nitrogen positive ions carry, very approximately, 95 per cent. of the current at the cathode, 80 per cent. one quarter of the way across the dark space, 55 per cent. one-half of the way across, and 25 per cent. about three-quarters of the way; these numbers are not inconsistent with a large positive wall charge close to the cathode, falling off at first slowly on approaching the anode side of the cathode dark space, then more rapidly, and finally changing to a negative charge at the boundary of the negative glow. The case of oxygen has not been dealt with by Compton and Morse, but there is no reason to suppose that the wall potential should be anomalous, except possibly at the edge of the negative glow [2], where the curves do in fact frequently show a rather large initial slope (dV/dx), as already mentioned. The results are also consistent with the fact, established in some work now in progress in this laboratory, that the current to a collector in the dark space is positive at the space-potential except close to the negative glow.

The curves of figs. 3–8 are similar to those for the potential assumed by an insulated collector which projects into the dark space, the interpretation of which has been discussed by Aston [7].

Another factor which would tend to increase the positive wall charge would be a secondary emission of electrons; nothing is known of the magnitude of this under the conditions of this work, but it is sometimes of great importance.

A variable type of discharge described by Aston [8] may be due to the wall charges in the dark space. In a glow discharge through helium at a definite pressure and current density, under conditions generally similar to those of the present investigation, the discharge assumed one of two forms erratically. The one was characterized by a low voltage and a diffuse edge to the negative glow, and the other by a high voltage and a sharper edge. A somewhat similar phenomenon in high vacuum has been described by Langmuir [9], where two distinct types of discharge can occur between a filament and a plate, the one accompanied by a large secondary emission from the walls and the other by a smaller secondary emission. It is possible that the two types of discharge in helium are due to the existence of two different arrangements of the wall charges, although they might also be connected with the presence or absence of oscillations. Aston suggested that they were connected with the purity of the gas. In some work with helium now in progress here a diffuse edge has invariably been found.

The effect of the wall charges upon the indirect measurements of ionization constants by Brown and Thomson's method is still difficult to estimate. The appearance of the feeble light from the cathode dark space suggests that there is present on the walls a somewhat indefinite sheath which is less bright than the body of the discharge, and that its thickness is greatest a short distance out from the cathode, and decreases to zero at the negative glow. Whilst transverse observation of a luminous cylinder, the intensity of the light from which was constant over a cross-section but diminished towards the cathode, would give rise to a similar appearance because of the less thickness of gas viewed nearer the wall, there is some independent evidence for a repulsion of the discharge from the walls in Childs' [10] observation that the current density is not uniform over the cathode, even with abnormal fall in potential. This could be due to a focussing of the current in an effective volume converging towards the cathode, similar to that observed by Oliphant and Moon [11] near the edges of a flat cathode remote from the walls. The

sheath, if present, is probably usually not very thick, for Brown and Thomson found that all parts of their collector were in fair adjustment to the space potential simultaneously, as judged by the shadows thrown. It seems likely, all things considered, that the effect of the walls in Brown and Thomson's work was not a greater cause of uncertainty than incomplete knowledge of the effect of the dissociation of polyatomic gases, and of the exact interpretation to be put on the constants in Morse's equation ; it is probable that the difference between their values for the ionization constants and those given by Townsend are due both to these causes and to a real difference between the ionization constants in Townsend's experiments with gas at high pressure passing small currents, and in Brown and Thomson's tube with more highly ionized gas at low pressure.

The bearing of our results on the failure of discharge-tubes is discussed elsewhere by R. H. Sloane *.

Summary.

The potential of the walls in the cathode dark space of a glow discharge has been measured by an electrostatic method for cathode falls in potential of not greater than 440 volts. The walls are strongly positive to the discharge close to the cathode and negative in the negative glow, so that walls and discharge are at the same potential in one place, which proves to be at the boundary of the negative glow and cathode dark space. The bearing of the results obtained on the measurements of potential made in the dark space by Brown and Thomson is discussed briefly, and the difficulty of obtaining a theory of the wall potential pointed out.

References.

(1) Brown and Thomson, Phil. Mag. viii. p. 918 (1929).
(2) Carmichael, Phil. Mag. viii. p. 362 (1929).
(3) Morse, Phys. Rev. xxxi. p. 1003 (1928).
(4) Carmichael and Emeléus, Phil. Mag. viii. p. 909 (1929).
(5) Compton and Langmuir, Rev. of Modern Physics, ii. no. 2 (1930).
(6) Compton and Morse, Phys. Rev. xxx. p. 305 (1927).
(7) Aston, Proc. Roy. Soc. lxxxiv. p. 526 (1910).
(8) Aston and Watson, Proc. Roy. Soc. lxxxvi. p. 168 (1912).
(9) Langmuir, Gen. Elect. Rev. xxiii. p. 513 (1920).
(10) Childs, Phil. Mag. ix. p. 529 (1930).
(11) Oliphant and Moon, Proc. Camb. Phil. Soc. xxv. p. 461 (1929).

* Phil. Mag., Dec. 1930, p. 1099.

III. *A Microscopic Method.* *By* E. H. Synge *.

SOME time ago the present writer suggested in the 'Philosophical Magazine' a method by which microscopic resolution might be considerably increased. This method consisted essentially in investigating the opacity of a microscopic section by means of a spot of intense light whose dimensions were small in comparison with a wave-length. The section was to be moved across the spot of light, and a picture built up, as in telephotography, by means of photoelectric apparatus. The form in which the writer endeavoured to develop this idea was, however, open to serious objections. He proposed to obtain the spot of light by utilizing a very minute hole, such as occurs in the silvering on glass, but this made it necessary that the section should be passed over the hole at a constant distance from the silvered plate, very small in comparison with a wave-length, and also that the surface of the section should be truly plane, and, of course, perfectly clean.

The technical difficulties involved in these conditions became more apparent on further consideration, and he also found that he had taken too high a figure for the opacity of silver, and hence that the possible increase in resolving power, although still considerable in theory, would not be so great as he had supposed, even if the formidable technical difficulties were overcome.

He was thus led to consider the possibility of utilizing the idea in some quite different form. It appeared that if the spot of light could be *projected* into the interior of the microscopic section the most serious of the technical difficulties would disappear. The problem, in fact, reduced itself to that of finding some practicable means by which a spot of intense light, whose dimensions were small in comparison with a wave-length, could be projected into the interior of a microscopic slide. The smallest spot of light which can be projected by an objective lens used as a condenser has, of course, the same limits as the microscopic resolution itself: so that the solution of the problem does not lie in this direction. Nor did anything in the nature of a diffraction screen, used in conjunction with such a condenser, seem to afford a means of projecting a

* Communicated by the Author.

spot of light of the kind required—a very small region of great intensity surrounded by a large region of small intensity. It became clear, indeed, that the problem was insoluble so long as the approximate theory of diffraction remained binding. This theory is, in most cases, a very good approximation, and forms the basis of the theory of resolution of optical instruments as usually presented, but it is by no means an absolute theory, such a theory requiring the solution of the electromagnetic equations, subject to boundary conditions which have been a bar to their solution except in one very simple case. In general we may say that when we come down to magnitudes of the order of a wave-length the approximate theory ceases to be a good approximation. It is pointed out, for instance, in Schuster and Nicholson's ' Optics ' that the total energy coming through a slit of such width is not correct according to the approximate theory, and one may fairly argue from analogy that in the case of a high-power microscope objective, resolving to about a wave-length, there is an appreciable difference between the true theoretical value of the resolution, if this could be determined, and the value derived from the approximate theory, which is often, but incorrectly, presented in text-books as if its validity and accuracy were not questionable in the case of a high numerical aperture. With the highest powered objective in use the angle of collection does not, however, much exceed 120°, and the difference between the true theory and the approximate theory is probably inconsiderable. But if we could proceed further, the difference would certainly become much more pronounced. If it were practicable to have an angle of collection of 180° the true theoretical value of the resolution might very probably be twice, or nearly twice, as high as the theoretical value on the approximate theory of diffraction.

At first sight an angle of 180° seems to present an absolute upper limit, even in theory, to the angle of collection (or projection). But after many unsuccessful attempts to solve the problem the writer finally saw how it would be possible to employ an angle of approximately 360°, and in so doing to free ourselves completely from the approximate theory of diffraction, and at the same time be provided by the absolute theory with a completely soluble problem which could be legitimately extended to the actual case. This method of using the full angle of 360° was not

indeed available for direct vision, but it was available for projecting a spot of intense light, very small in comparison with a wave-length, into the interior of a microscopic section, and it therefore met the requirements of the method of microscopic photography which he proposed.

We shall first consider the following general problem. A hollow cavity is constructed in the form of an ellipsoid of resolution, the walls of the cavity being silvered. A point source of light, *i.e.*, one negligibly small in comparison with a wave-length, is situated in one focus of the ellipsoid. An image of the point source will be formed in the other focus. What will be the size of the image? We cannot answer this question completely, but we can give an approximate answer for various cases.

In the first place, if we take the case of a very elongated ellipsoid it is clear that the bulk of the light which reaches the second focus comes from a small portion of the ellipsoidal surface near the first focus. It will therefore travel to the second focus in a beam converging at a small angle, and the small fraction of light travelling to the second focus from the remainder of the ellipsoidal surface cannot materially affect the phenomenon. Now in the case of such a narrow pencil the approximate theory of diffraction supplies us with a very good approximation, and we can predict with certainty that the diameter of the image measured in a plane perpendicular to the axis of the ellipsoid will be large in relation to a wave-length, and will tend to increase with the eccentricity of the ellipsoid.

Taking now a second case, in which the distance between the foci of the ellipsoid is equal to the latus rectum, the approximate theory of diffraction will still indicate what kind of result we may expect. The diameter of the image, measured in a plane perpendicular to the axis, will probably be about a wave-length—certainly of that order. We would, of course, define the image in some conventional way, as, for instance, the region where the intensity does not at any point fall below one-half of the maximum intensity in the image.

As the ellipsoid becomes more nearly a sphere the approximate theory of diffraction ceases to afford us any guidance, but in the extreme case, to which the others approximate—that of a sphere—we have a problem whose solution can be stated with perfect certainty. If a point source of light is situated at the centre of a sphere the

waves reflected from the surface of the sphere will return to the centre and form a point image of the source there. The size of the image will not bear any relation to the wave-length, and if the ellipsoid is very nearly a sphere we can argue with perfect confidence that the image will still be very small compared with a wave-length.

Lest the reader should have any doubt that the wave emitted from a point source of light at the centre of a sphere and then reflected by the surface will return again to form a point image at the centre, it may be well to consider one case in which a complete solution can be obtained without difficulty, and which is, in fact, equivalent to that case with which we shall actually be concerned.

This is the case of a Hertzian doublet, and to simplify the boundary conditions we will suppose that the reflexion takes place at a glass-air or air-glass surface, *i. e.*, the sphere is a cavity in a block of glass, or else, and this case is more nearly the actual one which we will soon consider, the sphere is a solid sphere of glass at whose centre is situated a Hertzian doublet, very small in comparison with a wave-length, which sends out spherical waves of light.

The mathematical problem is quite analogous to that of finding the reflected and refracted waves in the case of a plane wave incident on a plane surface. If, as in the usual treatment of the Hertzian doublet,

$$\Pi_1 = \frac{\phi_1}{r} \sin \frac{2\pi}{\tau}\left(t - \frac{r}{v_1}\right)$$

represent the original wave, v_1 being the speed of light within the sphere, we assume that the refracted and reflected waves are represented respectively by

$$\Pi_2 = \frac{\phi_2}{r} \sin \frac{2\pi}{\tau}\left(t - \frac{r}{v_2} + \epsilon_2\right),$$

$$\Pi_3 = \frac{\phi_3}{r} \sin \frac{2\pi}{\tau}\left(t + \frac{r}{v_1} + \epsilon_3\right),$$

where v_2 is the velocity of light outside the sphere and ϕ_2, ϵ_2, ϕ_3, ϵ_3 are constants to be determined. If the boundary conditions yield four equations to determine these four quantities, we conclude that our assumptions are justified.

Taking α_1, β_1, γ_1, X_1, Y_1, Z_1 as the components of magnetic and electric intensity respectively corresponding

to the original wave, and K_1 as the specific inductive capacity within the sphere, and assuming that the doublet acts along the Z axis, we have

$$\alpha_1 = \frac{\partial^2 \Pi_1}{\partial y \partial t}, \quad \beta_1 = -\frac{\partial^2 \Pi_1}{\partial x \partial t}, \quad \gamma_1 = 0,$$

$$X_1 = \frac{c}{K_1}\frac{\partial^2 \Pi_1}{\partial z \partial x}, \quad Y_1 = \frac{c}{K_1}\frac{\partial^2 \Pi_1}{\partial z \partial y}, \quad Z_1 = \frac{c}{K_1}\left(\frac{\partial^2 \Pi_1}{\partial x^2} + \frac{\partial^2 \Pi_1}{\partial y^2}\right).$$

There will be similar equations for Π_2 and Π_3, in the case of the former K_2, the specific inductive capacity outside the sphere, being substituted for K_1.

The boundary conditions may be written

(1) $\alpha_1 + \alpha_3 - \alpha_2 = 0$,

(2) $\beta_1 + \beta_3 - \beta_2 = 0$,

(3) $l(X_1 + X_3 - X_2) + m(Y_1 + Y_3 - Y_2) + n(Z_1 + Z_3 - Z_2) = 0$,

for all values of l, m, n satisfying the relation

$$lx + my + nz = 0.$$

These conditions must be satisfied all over the sphere and for all values of t. On transforming to polar coordinates it is immediately evident that the first two conditions become

$$\frac{\partial^2}{\partial r \partial t}(\Pi_1 + \Pi_3 - \Pi_2) = 0,$$

which can be written

$$A \sin\frac{2\pi t}{\tau} + B \cos\frac{2\pi t}{\tau},$$

where A and B are expressions independent of t, but involving the four undetermined constants ϕ_2, ε_2, ϕ_3, ε_3.

The third condition, after reduction, becomes

$$\frac{1}{r}\frac{\partial}{\partial r}\left(\frac{1}{K_1}\Pi_1 + \frac{1}{K_1}\Pi_3 - \frac{1}{K_2}\Pi_2\right) + \frac{4\pi^2}{c^2\tau^2}(\Pi_1 + \Pi_3 - \Pi_2) = 0,$$

which can be written

$$A' \sin\frac{2\pi t}{\tau} + B' \cos\frac{2\pi t}{\tau},$$

where A′ and B′ are expressions independent of t, but involving the four underermined constants ϕ_2, ε_2, ϕ_3, ε_3.

Now the necessary and sufficient conditions that

$$A \sin\frac{2\pi t}{\tau} + B \cos\frac{2\pi t}{\tau} \quad \text{and} \quad A' \sin\frac{2\pi t}{\tau} + B' \frac{\sin 2\pi t}{\tau}$$

may both be zero for all values of t, are

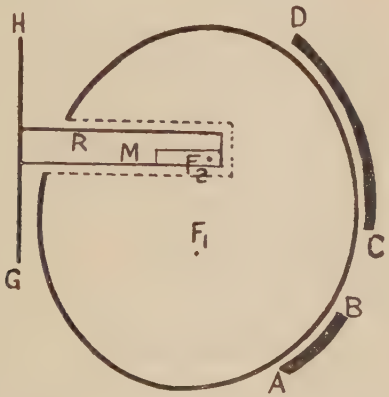
$$A = B = A' = B' = 0.$$

We thus have four equations for the four undetermined constants. The actual values can be found without great difficulty, and, putting in the values for ϕ_3 and ε_3, we have the equation of the reflected wave, showing us that this returns to the centre of the sphere, carrying back all the reflected energy into a region at the centre very small in comparison with a wave-length.

Now it is known, both by theory and experiment, that the light scattered from a particle which is small in comparison with a wave-length has equations of the same form as a Hertzian doublet. It will therefore be evident how we may obtain, within a solid ellipsoid of glass which is nearly a sphere, a point of intense light which is small in comparison with a wave-length. We need only suppose that a small colloidal particle of gold (of the order of 10^{-6} cm.) is situated at one focus, and that a beam of intense light illuminates the particle, passing out from the ellipsoid at a portion of the surface which is modified so that the light of the beam will not be reflected back into the ellipsoid. The light refracted from the colloidal particle will then spread outwards in waves which are nearly spherical, and will, in part, be reflected by the surface of the ellipsoid, except by that small portion which is modified to admit of the emergence of the illuminating beam. The reflected wave will travel inwards and will form an image at the other focus which will be small in comparison with a wave-length, although, no doubt, for various reasons, somewhat larger than the colloidal particle.

The reader may not see immediately how the power to produce such a spot of light in the interior of a solid ellipsoid of glass can be applied for our purposes to a microscopic section; but the diagram will probably make this clear. The eccentricity of the ellipse in the diagram is, of course, shown with great exaggeration. If drawn correctly F_1 and F_2 would scarcely appear to the eye as separate points.

The colloidal particle P is situated at the focus F_1. AB is a condensing lens by which intense light from a small source is brought to a focus at F_1. The dotted line represents a hole or well which is made in the ellipsoid, and R is a glass rod which moves freely in this hole, the intervening space between R and the walls of the hole being filled with immersion fluid having the same index of refraction as the glass of the ellipsoid and the rod. M is a small microscopic section which is attached to R, R being of course practically a narrow microscope slide. The imbedding medium of the section has the same index of refraction as the glass. HG is a frame in which R is fixed, HG being capable of receiving small regular displacements of the order of 10^{-6} cm. by means of levers or



some other device. The illuminating beam from AB emerges from the ellipsoid where R enters it. This economizes the surface. F_2, the second focus of the ellipsoid, is within M. CD is a collecting lens, so adapted that F_2 is in its focal plane, and rays emerging from F_2 are collected by CD and ultimately brought to a focus, as a diffraction disk, at a point p. The ellipsoid may with advantage be partially silvered, but this does not affect the general history of the rays.

Let us consider this history. We have a wave which is virtually spherical spreading outwards from F_1. Portion of this is reflected by the surface of the ellipsoid, and travels inwards to F_2, carrying all its energy into a region in M which is slightly larger than the particle P. This

assumes, of course, that the wave is not much disturbed by opaque matter or matter with a different refractive index as it travels through M to F_2, a condition which we would seek to satisfy, as far as possible, by making M very thin and very small. After passing through the focus at F_2 the wave will spread out once more, and a portion of it, passing through the surface of the ellipsoid, will be collected by CD and brought to a focus at p. The intensity of the light at p, as determined by some refined photoelectric apparatus, will, practically speaking, measure the opacity at the particular point of the microscopic section M which coincides with F_2. And if M be given regular displacements of the order of 10^{-6} cm. by means of the frame HG which carries it, different points of M being brought consecutively into coincidence with F_2, the opacity of all points in M throughout a small region can be determined, and, if desired, a photograph can be built up as in phototelegraphy. Light coming directly from F_1 will, of course, also be collected by CD, but in general this light will not be in correct focus, or, if it is, it will come to a focus at a different point than p. There will also be some light arriving at p after three reflexions at the surface of the ellipsoid, but this will be inconsiderable and will only have the effect of making the results a very little less distinct.

Our principal object being to make the image at F_2 as small as possible, it may be well to consider the various factors which tend to increase its size. In the first place there is the portion of the ellipsoid where the beam from AB emerges, and which does not contribute to the reflexion. This cannot be decreased indefinitely, since the intensity of the illumination at F_1 is measured by the solid angle formed by the beam converging at F_1, and we require an intense illumination. We have no simple theory to guide us as to the effect upon the size of the image at F_2, of the omission of a certain small portion of the surface from the reflexion. If, for instance, the wave is reflected by $\frac{99}{100}$ of the surface, we can only conjecture how much the size of the image will be increased by the default of the remaining $\frac{1}{100}$. Some experimental work, using the infra-red, would probably throw light on this. As it is

very important to have intense illumination, the angle of the illuminating cone should be widened until it is found that the size of the image begins to increase materially.

A second factor has already been touched on. The spherical wave moving inwards to its focus at F_2 will pass through the microscopic section in every direction, and will be disturbed more or less by any opacity or difference of refractive index anywhere in the section. Such disturbance will lead to an increase in size of the image ultimately formed at F_2. The only way to deal with this cause of disturbance is to make the section as small and thin as possible. There is no advantage, in any case, in having a large section, for the area which could be dealt with in a reasonable time by the method proposed is necessarily small—a few μ square. In one very important case such questions do not arise. Organisms of "ultra-microscopic" size, if sparsely enough distributed in a medium of the correct refractive index, should reveal their shapes and the relative opacity of their different parts to a limit of resolution, corresponding very nearly to the size of the image which would be formed at F_2. if the section were perfectly transparent. It should not be necessary to trouble about limiting the size of the region through which they were distributed. Their presence and approximate position would be revealed to the eye by faint diffraction disks and HG adjusted to bring them to F_2. For this important case the apparatus might be more than a laboratory instrument.

Another factor affecting the size of the image at F_2 is the correctness of the surface of the ellipsoid. To obtain a resolution approaching ten times that at present attainable, the surface should be correct to beyond 10^{-6} cm. The writer is not certain as to the exact limit of accuracy attainable in the figuring of a glass sphere by methods at present in use, but it would probably be inadequate for so high a resolution. The principle of interference, in conjunction with photoelectric appliances, is, however, capable of supplying tests for accurate figuring up to and beyond our requirements. If we have two surfaces parallel to one another, in the case of planes, or concentric in the case of spheres, and if a beam of monochromatic light strikes the two surfaces normally, the monochromatic light being chosen of such a wave-length that the two reflected beams have a difference of phase of about one

quarter, a variation as small as $\frac{1}{2000}$ to $\frac{1}{5000}$ of a wavelength in the distance between the two surfaces should produce an appreciable variation in the reading of a photoelectric cell, employed to measure the intensity of the combined reflected beams. In the case of spherical surfaces two surfaces would be prepared together—a convex sphere and a somewhat larger concave hemispherical shell. These would be adjusted to concentricity by means of springs and a beam of monochromatic light passed through a lens L, so as to converge to their common centre. The rays of this beam would, of course, strike the two surfaces normally. There should be connexion between L and the outer surface of the hemispherical shell through immersion fluid, and the half of the sphere which is not under immediate investigation should be placed in similar fluid of the correct refractive index. Complicated reflexions would thus be avoided. The beam should enter L originally as a parallel beam, and before entering L should pass through a half-silvered plate at 45°. The two reflected beams, which we assume to have a phase difference of about one quarter, will pass through L in the reverse direction, and, issuing as parallel beams, be in part diverted by the half-silvered plate. This part will be received by a photoelectric cell, which will register the combined intensity. As different portions of the two surfaces are brought in turn under review the reading of the cell should remain unchanged if the surfaces are accurate and accurately centred (inaccurate centreing can easily be recognized and allowed for). It should also remain unchanged if they are turned about relatively to one another in any way. An instrument designed on these lines should make it possible to test spherical surfaces to beyond 10^{-7} cm., which considerably exceeds the limit required.

As regards the passage from spherical to ellipsoidal surfaces, it should be noted that the modification of a sphere which is required is exceedingly small. If, for instance, the sphere was 2 cm. in diameter, a maximum surface modification of ·0005 mm. would give an ellipsoid with a distance between the foci of about 2 mm., which would allow of the use of a thin cover-glass over the section. The modification of a sphere to this extent should present no difficulty if the rate at which a polishing

machine removed the surface were known with some accuracy.

Of all the factors which would increase the size of the image at F_2 the one most likely to escape proper attention and most difficult to provide against is probably the difference between the refractive index of the glass and that of the fluid filling the space between R and the walls of the slot; for if this fluid has not the same dispersion as the glass the two refractive indices will not coincide except at one point of the spectrum. And even if the refraction and dispersion are correct at one temperature, they will not be correct at another, since the refractive index in all fluids changes more rapidly with temperature than in glass. To make any difference in the refractive index as harmless as possible the sides of R should be parallel to the faces of the slot and as near them as possible, no more immersion fluid being used than is needed to allow R to move freely. The precise degree of parallelism would depend on the difference of refractive indices which must be reckoned with, and should be calculated for the case. Probably it would be possible to obtain a mixed fluid, or solution, in which the refractive index was practically the same as the glass throughout the spectrum at a certain temperature, and in preliminary experiments it would be most satisfactory to work consistently at this temperature, and so eliminate this factor completely. If we define the image at F_2 as the region within which the intensity is nowhere less than half the maximum intensity attained in the neighbourhood of F_2, it is probable that if all care were taken the diameter of the image should not be much greater than that of the particle at F_1, which will then represent the resolving power of the apparatus.

It should be remarked that, for preliminary experiments at all events, an unmodified sphere might be used, dispensing with a cover-glass. In the case of a sphere any two points at equal, very small, distances from, and at opposite sides of, the centre are mutual foci. There is therefore no need for accuracy in placing P with respect to the spherical surface. The best method of ensuring that P was properly placed with relation to the surface would, in any case, probably be to figure the sphere or ellipsoid from a piece of glass containing a very few scattered particles of gold, one of them of the size desired (as indicated by its brightness) being selected, and the

sphere or ellipsoid being *then* figured so that this particle occupied, as accurately as possible, the proper position within it. The wall of the slot would, in general, come within a few μ of the particle. From experiments by Zsigmondy*, in which he seems to have used a numerical aperture of $\frac{1}{3}$, with sunlight, it appears that, with such a concentration of light, which is as great as would be employed in the present case, the substance of ruby glass, even with gold amicrons present, did not scatter light to any appreciable extent. Thus one might take a minute fragment of ruby glass, containing a few gold particles of suitable brightness, and remelt it at not too high a temperature in a large bulk of glass (without gold) of the same composition. The resulting glass should have a few particles of gold widely scattered through it. One of these would be selected in the manner proposed, the occurrence of a few other particles throughout the ellipsoid, far away from the focus, being of no consequence. The approximate sphericity of the selected particle, which is desirable, could probably be estimated by illuminating the particle within the block of glass from different directions, the observer, who would use a large lens like a condenser lens, such as CD in the diagram, preserving his direction with respect to the illuminating apparatus. If the particle were not a sphere it would presumably show different effects when illuminated from different directions.

It is a technical point, which need not be considered here, whether it would not be more convenient for figuring the slot, to make the sphere in two separate hemispheres, cementing these together before the final figuring of the surface was carried out. The cement would, of course, introduce a new factor which would have to be reckoned with.

We have still to deal with a question which is vital to the proposed method, and in which there is some difficulty in drawing a conclusion from the authorities. Will a gold particle of about 10^{-6} cm. in diameter scatter sufficient light (using sun or arc illumination, with a numerical aperture of $\frac{1}{3}$ or $\frac{1}{4}$) to operate a photoelectric recording apparatus ? If one were to approach the question in a quite theoretical fashion the answer would seem to

* Zsigmondy, 'Colloids and the Ultra-microscope,' p. 174 (invisibility of light-cone with small slit).

be in the affirmative. First we have Rayleigh's law for very small particles, according to which the amount of light scattered by such a particle (assumed spherical) varies as the *sixth power* of the diameter *. This law could only be expected to hold while the particles are so small that an inconsiderable fraction of the light is absorbed in passing through them. When this stage is passed—at about 10^{-6} cm.—similar reasoning to that employed for Rayleigh's law would point to the law becoming one of variation as the *fourth power* of the diameter. When a further stage is reached and the diameter becomes of the order of a wave-length the scattering assumes the character of ordinary reflexion, and the law becomes one of the *second power*. Basing oneself on these considerations, and calculating backwards for a roughly approximate result, it would appear that a gold particle of about 2×10^{-7} cm. in diameter should represent the limit of visibility under the strongest concentration of sunlight attainable in an ultra-microscope. This is very close to Zsigmondy's figure $1{\cdot}7 \times 10^{-7}$ cm.†, and when one considers the roughness of such a *per saltum* approximation there seems little doubt that the general contention is correct. The assumptions are indeed based on good grounds, and in the case of Rayleigh's law for very small particles are supported by a mass of experimental evidence. Indeed, Rayleigh's law does not seem to be challenged by any authority, and if it holds for particles up to 10^{-6} cm. in diameter, and if we accept Zsigmondy's figure $1{\cdot}7 \times 10^{-7}$ cm. for the minimum of visibility, it follows that a particle 10^{-6} cm. in diameter should send to the eye, under similar illumination etc., 40,000 times the minimum visible quantity of light. With arc illumination and a numerical aperture in the condenser of about $\frac{1}{3}$ this would reduce to about 1000.

It would be unnecessary to discuss the matter further were it not that, side by side with Rayleigh's law and Zsigmondy's figure, other figures are given which are entirely irreconcilable with them. Burton, for instance, again quoting Zsigmondy, gives $1{\cdot}5 \times 10^{-6}$ cm. as the diameter of the smallest particle of gold visible by arc

* Burton, 'The Physical Properties of Colloidal Solutions,' pp. 105, 110 (1921).

† Burton, *op. cit.* p. 48.

illumination *. A simple calculation, however, shows that on the basis of Rayleigh's law and the ordinary estimate of the comparative brightness of sun and arc-light, such a particle should send to the eye no less than 100,000 times the minimum visible quantity of light, assuming that Zsigmondy's other figure is correct. The reader will, no doubt, share the present writer's surprise that this enormous discrepancy is not even commented on. Rayleigh's law is enunciated without question, and the figures are given equally without question.

The source of the discrepancy is, however, not difficult to discover; for the ordinary method of estimating the size of a colloidal particle, as is repeatedly pointed out by Zsigmondy, gives simply *an upper limit* to the average size. The calculated size is, in fact, a function of the brightness of the illuminating source, of the condensing system, and of the dilution, and varies largely as these vary. The two figures of Zsigmondy seem to have been obtained with quite different condensing systems, and, apart from this, both represent upper limits only. It appears to the present writer that the authorities on the subject, even Zsigmondy, do not keep consistently in view the dependence of their calculated values on these factors, and that in the result their figures, given generally without specification of the illuminating system (without which they are really meaningless), are, as a rule, several times too great if taken to represent the actual size of the particles. If we accept Zsigmondy's figure $1{\cdot}7 \times 10^{-7}$ cm. as an upper limit (which we seem justified in doing) to the diameter of the smallest visible particle of gold, then a particle 10^{-6} cm. in diameter *must* appear in the ultra-microscope, even with arc illumination, as a brilliant object. Otherwise we should have to reject Rayleigh's law *in toto*, although it is inseparable from theory which has received abundant experimental confirmation. This means in fact that often particles estimated by the ordinary method as having diameters of perhaps 4×10^{-6} cm. or 3×10^{-6} cm. are really no more than 10^{-6} cm. in diameter, a conclusion which is borne out by Svedberg's results on the Brownian movements†, and by many other facts and anomalies cited in

* *Op. cit.* p. 122.

† Burton, *op. cit.* p. 76.

Burton's monograph. On the other hand, the writer has failed to find a single fact opposed to this conclusion.

It seems, therefore, that we may presume with confidence that a particle 10^{-6} cm. in diameter will scatter sufficient light to operate a photoelectric cell, the current being of course amplified for recording purposes. A method could certainly be suggested, by which selenium could be employed to record automatically much smaller amounts of light than those posited, in brief intervals of time, and in which the result would not be affected by the lag and irregularities of the law governing selenium's response to light. At the same time everything should be done to make as much light as possible available. With this object the surface of the ellipsoid should be partially silvered, so as to reflect about 25 per cent. of the light. The silvering should be very even over the surface, and this might be easier to effect if some metal less opaque than silver were used. In the case of an ellipsoid coated in this way with metal it would probably be of advantage to keep the entire ellipsoid in immersion fluid, the connexion with the lenses AB and CD being through this fluid. It would then be unnecessary to modify the faces of these lenses. The amount of light available could also be increased by placing collecting lenses of the type of CD all round the ellipsoid, thus collecting the light which issued from F_2 in every direction. Each lens would bring its share of this light to a focus, but the beams being diverted by mirrors in their course all these foci would be brought to bear upon the same photoelectric surface, the light coming directly from F_1 being, of course, cut out as far as possible. Thus the greater part of the light issuing from F_2 would contribute to the current in the cell.

A word may be added as to the small even displacements which would be required of the frame HG. These would correspond with the resolution aimed at, say 10^{-5} cm. to 10^{-6} cm. The best way of obtaining such displacements would probably be either by levers or by wedges. The latter seem especially suitable, and it is difficult to understand why refined wedges are not generally used when very small displacements over a very small range are required. They would have every advantage over screws (which are essentially only spiral wedges), for, however carefully a

screw may be ground, it is impossible that the face of its thread should be optically accurate. On the other hand, wedges could be constructed whose faces were optically plane, and which made an angle of as little as a few seconds of arc, if desired, with one another. Such wedges would seem to offer a means of making the smallest regular displacements over a restricted range. They would, of, course, be actuated by screws, and used in conjunction with other screws which would make the coarser displacements.

Most of the points have now been discussed which would probably present themseles to anyone proposing to construct and operate an actual instrument on the lines suggested. Some of them would hardly arise if the experimenter limited himself at first to the attempt to construct an instrument which would give a resolution about three times beyond that attainable by ordinary microscopic methods. For this case an unmodified sphere might be used, constructed as accurately as was practicable by the usual methods. One might safely employ an illuminating beam of fairly wide angle, say 60° to 90°, and there should be abundance of light at p, even if the sphere were not silvered, to give good quantitative results, especially if one used a fairly large colloidal particle, at F_1. The displacements of HG which would be required in this case should not be smaller than could be attained by means of very fine screws.

The writer feels so little doubt as to the practicability of an instrument of the kind, which, with certain technical refinements of a fairly simple character, should give an increase of resolution up to 10 or 20 times beyond the ordinary microscope, that he will venture to give a name to it, and to call it a *hypermicroscope*. The technical refinements and the inconvenience, which might prove unavoidable, of having to keep portion of the apparatus at a fixed temperature appear inconsiderable in comparison with the importance to all the biological sciences which such an increase of resolution should have ; while if the form of the proposed apparatus seems unfamiliar, and even a little bizarre, this is only because it embodies a point of theory which has not been utilized before, but which probably offers the only means by which light waves can be made to yield a resolution which is a very small fraction of a wave-length.

IV. *The Diamagnetism of Polyhalides.* *By* Francis W. Gray, *M.A., D.Sc., Senior Lecturer in Chemistry, Aberdeen University, and* John Dakers, *B.Sc.* *

THE following values for molecular diamagnetisms were determined by Dakers in this laboratory.

The figure in brackets after each numerical result indicates the number of measurements. For example, $87{\cdot}26 \pm 0{\cdot}46$ (6) means that six measurements gave an average value of 87·26, with an average deviation from the mean amounting to $\pm 0{\cdot}46$.

Table I.

Monohalides.

$N(CH_3)_4I$. $105 \pm 0{\cdot}88$ (3)	$N(C_2H_5)_4I$. $152{\cdot}6 \pm 0{\cdot}31$ (3)	$N[C_6H_5(CH_3)_3]I$. $144 \pm 3{\cdot}2$ (4)
$N(CH_3)_4Br$. $87{\cdot}21 \pm 0{\cdot}46$ (6)		$N[C_6H_5(CH_3)_3]Br$. 132·8
C_6H_5Cl. $69{\cdot}61 \pm 0{\cdot}35$ (4)	RbCl. 45·04 (1)	
	RbBr. $56{\cdot}69 \pm 0{\cdot}54$ (4)	
C_6H_5I. $90{\cdot}16 \pm 0{\cdot}51$ (4)	RbI. $67{\cdot}05 \pm 1{\cdot}9$ (4)	

Polyhalides.

Negative ions. \ Positive ions.	$N(CH_3)_4$.	$N(C_2H_5)_4$.	$N[C_6H_5(CH_3)_3]$.	C_6H_5.	Rb.
I_3	$181{\cdot}9 \pm 1{\cdot}6$ (3)	$224{\cdot}9 \pm 2{\cdot}9$ (5)	$211{\cdot}2 \pm 5{\cdot}9$ (6)		
I_2Br	$161{\cdot}3 \pm 0{\cdot}89$ (3)		$199{\cdot}1 \pm 0{\cdot}8$ (2)		
IBr_2	$150{\cdot}8 \pm 0{\cdot}78$ (4)	$202{\cdot}7 \pm 0{\cdot}87$ (9)	$179{\cdot}6 \pm 1{\cdot}9$ (3)		$116{\cdot}2 \pm 0{\cdot}53$ (2)
IBrCl			$158{\cdot}0 \pm 3{\cdot}9$ (2)		
ICl_2			$166{\cdot}3 \pm 2{\cdot}7$ (3)	$117{\cdot}6 \pm 1{\cdot}2$ (5)	
Br_3	$137{\cdot}0 \pm 2{\cdot}1$ (4)		172·9 (1)		
$BrCl_2$	114·1 (1)		$156{\cdot}8 \pm 1{\cdot}6$ (2)		
I_5	263·1 (1)		$270{\cdot}8 \pm 4{\cdot}3$ (3)		

Values for ICl, $54{\cdot}54 \pm 0{\cdot}24$ (3), and for ICl_3, $90{\cdot}16 \pm 0{\cdot}69$ (4), were determined in this laboratory by Farquharson.

* Communicated by the Authors.

These diamagnetisms are discussed here by the method described in "Diamagnetism and Sub-Molecular Structure," Gray and Farquharson, Phil. Mag. x. pp. 191–216 (1930).

With the polyhalide anions in general it is clear that one of the halogen atoms has a special function in the constitution. It seems to act as a kernel round which the other halogen atoms are clustered. There is only one exception—perchloride of tetraethylammonium (Chattaway and Hoyle, J. C. S. cxxiii. p. 658, 1923)—to the rule that this kernel must be either an iodine or a bromine atom. With the one exception mentioned, every polyhalide anion has, for at least one of its atoms, either iodine or bromine, and there are more cases with iodien than with bromine. The capacity to act as a kernel would appear to increase with the tendency to positiveness, the order being as follows: $Cl < Br < I$.

We can account for the charge on the complex anion by assigning to each clustered halogen atom the negative valency 1, and to the kernel atom one of the positive valencies 1, 3, 5, and 7, the corresponding coordination numbers being 2, 4, 6, and 8.

In this way we obtain neutral groups such as

$$M^{+1}+I^{+1}+2I^{-1},\ M^{+1}+I^{+3}+4I^{-1},\ M^{+1}+I^{+5}+6I^{-1},$$
$$M^{+1}+I^{+7}+8I^{-1}.$$

Now when we compare the theoretical diamagnetism of the neutral group obtained in this way with the experimental diamagnetism of the molecule, we find that the latter is less than the former by so much that the most likely explanation is the development of homopolar bonds in which there is fluctuation of angular momentum causing feeble paramagnetism. For example, the neutral group $Rb^{+1}+I^{+1}+2Br^{-1}$ has the diamagnetism 200·27 and $RbIBr_2$ has the experimental molecular susceptibility 116·2. Thus the depression is 84·07. Again, with $C_6H_5ICl_2$ the neutral group has the diamagnetism 162 27, the experimental molecular diamagnetism is 117·6, and the depression is 44·67.

These depressions are accounted for by the following changes:—

	Depression below neutral groups.
$Rb^{+1}Br^{-1}I^{+1}Br^{-1} \rightarrow Rb^{+1}-Br^{+3}-I^{+1}-Br^{+1}$.....	84·07
$Ph^{+1}\,C^{-1}I^{+1}Cl^{-1} \rightarrow Ph^{+1}-Cl^{+3}-I^{+1}-Cl^{+1}$.....	44·67

Magnetism of Bonds.

	Three bonds.	Base-halogen bond.	Two halogen-halogen bonds.
$Rb^{+1}-Br^{+3}-I^{+1}-Br^{+1}$...	+34·97	$+(14·12-x)$	$+(20·85+x)$
$Ph^{+1}-Cl^{+3}-I^{+1}-Cl^{+1}$...	+15·78	$-(4·42+y)$	$+(20·20+y)$

The magnetism of the bonds in $Rb^{+1}-Br^{+1}$ and $Ph^{+1}-Cl^{+1}$ are +14·12 and −4·42 respectively. These bonds change on entering the trihalide, becoming less paramagnetic and more diamagnetic. They, however, probably retain their relation with regard to polar nature. Further, we expect the iodine-bromine bonds to be more homopolar than the iodine-chlorine, *i. e.*, $20·85+x>20·20+y$ or $x>y=0·65$ Also x should be greater than y because the change $Br^{+1}\rightarrow Br^{+3}$ should have a greater heteropolarizing effect than the change $Cl^{+1}\rightarrow Cl^{+3}$.

Accordingly each of the three bonds in the rubidium compound is probably more homopolar than the corresponding bond in the phenyl compound.

Polyhalides in general.

Each of the three cores Cl^{+7}, Br^{+7}, I^{+7} consists of complete groups round which a complete octet can be formed. For any halogen atom X, X^{-1} is X^{+7} surrounded by an octet, X^{+1} is X^{+7} surrounded by six electrons, X^{+3} is X^{+7} surrounded by four electrons. $-X^{+1}$ is X^{+1} with a bond of two electrons attached, and may change to X^{-1}. Similarly $-X^{+3}$ is X^{+3} with a bond of two electrons attached, and may change to X^{+1}. A change from halide to trihalide—for example, $KX\rightarrow KX_3$—may be represented thus:

$$K^{+1}-X^{+1}\rightarrow K^{+1}-\underset{a}{X^{+3}}-\underset{b}{X^{+1}}-\underset{a}{X^{+1}},$$

a pair of electrons being in each of the three bonds of the trihalide. The pair in the middle bond (b) is shared by two octets, but not those in the end bonds (a). The middle bond is, therefore, more homopolar than the end ones. Ionization takes place most readily at the end bonds; when it occurs at the base-halogen bond the complex trihalogen anion is obtained, thus:

$$K^{+1}-X^{+3}-X^{+1}-X^{+1}\rightarrow K^{+1}+(X^{+1}-X^{+1}-X^{+1})$$

or $\quad K^{+1}+(X_3)^{-1}.$

Similarly we could regard the higher polyhalides as having structures with bonds alternate in their polar nature. For example, the pentahalide is

$$KX_5 \quad \text{or} \quad \begin{array}{c} X^{+} \\ | \\ a \\ | \\ K^{+1}-a-X^{+3}-b-X^{+3}-b-X^{+1} \\ | \\ a \\ | \\ X^{+1} \end{array}$$

The more homopolar bonds, containing shared pairs, are at *b* ; the less homopolar bonds, not containing shared pairs, are at *a*.

Corresponding structures can be given for KX_7 and KX_9.

It is impossible to devise strictly analogous structures for the complex anions of the halogenates and the polyhalides.

All the internal bonds of the halogenate complex anion are alike in having a pair of electrons shared by two octets, and in displaying similar homopolar nature. Also ionization is equally difficult at all these bonds.

On the other hand, in the polyhalide complex anion there are two distinct kinds of internal bonds. The first class contain shared electrons, the second do not. At the first, ionization is much more difficult than at the second, and the nature is much more homopolar. The two types of bonds are indicated by the letters *b* and *a* respectively in the following structures :—

$$\begin{array}{c} O^{0} \\ | \\ b \\ | \\ K^{+1}-a-O^{+2}-b-I^{+5} \\ | \\ b \\ | \\ O^{0} \end{array} \qquad Rb^{+1}-a-Br^{+3}-b-I^{+1}-a-Br^{+1} \qquad \begin{array}{c} Cl^{+1} \\ | \\ a \\ | \\ K^{+1}-a-Cl^{+3}-b-I^{+3}-b-Cl^{+1} \\ | \\ a \\ | \\ Cl^{+1} \end{array}$$

The bond outside the complex anion, in both the halogenate and the polyhalide, is of the type *a*, heteropolar, the base-oxygen bond on the one hand and the base-halogen bond on the other.

The remarkable depression of diamagnetism observed when oxygen is added to a halide to produce a halogenate has its analogue in the depression occurring when halogen is added to a halide to form a trihalide of the polyhalide type, *i. e.*, when the addition is without valency change in the basic atom. Although the structures of the halogenate

and the trihalide are not strictly analogous, the explanation of the depression in both cases is to be found in the casting-off of electrons from certain of the original atoms and in the development of a feeble bond paramagnetism.

TABLE II.

Comparison of the same Trihalide of different Bases.

The change from Halide to Trihalide.

	Fall of diamagnetism.		
	Total fall.	Fall in cores. $X^{+1} \rightarrow X^{+3}$.	Fall in bonds.
$B^{+1}-I^{+1} + I^{+1}-I^{+1} \rightarrow B^{+1}-I^{+3}-I^{+1}-I^{+1}$			
B			
NMe_4	11·8	16·74	−4·94
NEt_4	16·5	16·74	−0·24
$NPhMe_3$	21·6	16·74	4·86
$B^{+1}-Br^{+1} + Br^{+1}-Br^{+1} \rightarrow B^{+1}-Br^{+3}-Br^{+1}-Br^{+1}$			
B			
NMe_4	11·41	12·72	−1·31
$NPhMe_3$	21·1	12·72	8·38
$B^{+1}-Br^{+1} + I^{+1}-I^{+1} \rightarrow B^{+1}-Br^{+3}-I^{+1}-I^{+1}$			
B			
NMe_4	14·7	12·72	1·98
$NPhMe_3$	22·5	12·72	9·78
$Cl^{+1}-Cl^{+1} + I^{+1}-Cl^{+1} \rightarrow Cl^{+1}-Cl^{+3}-I^{+1}-Cl^{+1}$	4·58	7·27	−2·69
$Ph^{+1}-Cl^{+1} + I^{+1}-Cl^{+1} \rightarrow Ph^{+1}-Cl^{+3}-I^{+1}-I^{+1}$	6·55	7·27	−0·72

The negative sign in the above table indicates that there is a rise of diamagnetism, instead of a fall, contributed by the bonds when the trihalide is formed.

The total fall of diamagnetism displayed in the reaction is, in several cases, nearly equal to the loss produced in the core diamagnetisms by the casting off of two electrons when one of the halogen atoms changes its charge from +1 to +3. However, changes in the magnetism of the bonds must also be reckoned with. The bond change is of two-fold origin. The raising of the charge of the halogen of the base-halogen

bond causes a rise of diamagnetism, and, on the other hand, the paramagnetism of the new bond (the homopolar middle bond of the trihalide) causes a fall. These two bond changes thus act oppositely, and may nearly cancel one another, or the bond change may be a fall or a rise of diamagnetism.

With the same trihalide the difference between different bases occurs mainly in the base-halogen bond. The raising of the charge of the halogen of the base-halogen bond has the effect of rendering this bond less homopolar or more diamagnetic. The substitution of methyl in the ammonium monohalides has the effect of raising the diamagnetism, and this rise may be associated with a stabilizing action in which certain bonds are rendered more heteropolar. If we suppose that this action increases with different substituents in the direction $NMe_4 < NEt_4 < NPhMe_3$, then there is more stabilizing (heteropolarizing) left to be done in the NMe_4-halogen bond than in the NEt_4-halogen bond, and more in that than in the bond with $NPhMe_3$. Accordingly, whereas the change in the base-halogen bond is usually predominant in the NMe_4 compounds, the contribution ot the middle bond has the greater influence in the $NPhMe_3$ compounds.

It is worth noting that for the three trihalides under comparison the depression of the diamagnetism of the bonds is in all three cases greater for the compound of $NPhMe_3$ than for that of NMe_4, and by an amount of the same order of magnitude in each case, being 9·80 in the triiodide, 9·69 in the tribromide, and 7·80 in the bromide-diiodide, or nearly equal in the first two instances.

On general grounds we prefer to regard the bond in $Ph^{+1}-Cl^{+1}$ as more heteropolar than that in $Cl^{+1}-Cl^{+1}$. This assumption is not without support from experimental results for the susceptibilities, and is in accordance with the fact that the contribution of the middle bond has much greater influence in $PhClI_2$ than in ICl_3.

The diamagnetic fall of 21·6 observed in the formation of the triiodide of $NPhMe_3$ is accounted for by a fall of 16·74 in the cores and 4·86 in the bonds. In the formation of the pentaiodide the loss is more than double that in the triiodide, being 50·8, accounted for by a fall of 33·48 ($2 \times 16{\cdot}74$) in the cores and 17·32 in the bonds. This excessive fall in the bonds agrees with the structure suggested above for the pentahalide.

On the other hand, the result for the pentaiodide of NMe_4 does not agree with our theory, but no importance can be attached to the figures for this compound, as only one measurement was made.

The negative sign in the table below indicates that there is a rise of diamagnetism (instead of a fall) caused by the change in the bonds in a step down the table.

TABLE III.

Comparison of different Trihalides of the same Base.

	Molecular diamagnetism.	Fall in diamagnetism in a step down the table.		
		Total fall.	Fall in cores.	Fall in bonds.
$B=NMe_4$				
$B^{+1}-I^{+3}-I^{+1}-I^{+1}$	181·9			
		20·6	21·46	− 0·86
$B^{+1}-I^{+3}-I^{+1}-Br^{+1}$	161·3			
		10·5	17·44	− 6·94
$B^{+1}-Br^{+3}-I^{+1}-Br^{+1}$	150·8			
		13·8	21·46	− 7·66
$B^{+1}-Br^{+3}-Br^{+1}-Br^{+1}$	137·0			
		22·9	29·79	− 6·89
$B^{+1}-Cl^{+3}-Br^{+1}-Cl^{+1}$	114·1			
$B=NEt_4$				
$B^{+1}-I^{+3}-I^{+1}-I^{+1}$	224·9			
		22·2	38·9	−16·7
$B^{+1}-Br^{+3}-I^{+1}-Br^{+1}$	202·7			
$B=NPhMe_3$				
$B^{+1}-I^{+3}-I^{+1}-I^{+1}$	211·2			
		12·1	17·44	− 5·34
$B^{+1}-Br^{+3}-I^{+1}-I^{+1}$	199·1			
		19·5	21·46	− 1·96
$B^{+1}-Br^{+3}-I^{+1}-Br^{+1}$	179·6			
		13·3	29·79	−16·49
$B^{+1}-Cl^{+3}-I^{+1}-Cl^{+1}$	166·3			
$B^{+1}-Br^{+3}-Br^{+1}-Br^{+1}$	172·9			
		16·1	29·79	−13·69
$B^{+1}-Cl^{+3}-Br^{+1}-Cl^{+1}$	156·8			

The bond change at each step thus yields a rise of diamagnetism, and this implies an increase in the heteropolar nature of the bond. For example, in the step $NMe_4I_3 \rightarrow NMe_4I_2Br$, the bonds become less homopolar, and we expect the iodine-iodine bond to be more homopolar than the iodine-bromine. The same agreement with theory is obtained throughout the whole table with the single exception of the step $NMe_4IBr_2 \rightarrow NMe_4Br_3$.

It should be noted that, in order to obtain this agreement with theory, NMe_4I_2Br has been written $B^{+1}-I^{+3}-I^{+1}-Br^{+1}$, whereas the corresponding $NPhMe_3$ compound is taken to be $B^{+1}-Br^{+3}-I^{+1}-I^{+1}$.

Experimental.

The Magnetic Measurements.

All the results were obtained with a Curie–Chéneveau magnetic torsion balance, with magnetic damping, circular movement of the permanent-magnet field, and reading of deflexion by means of a beam of light reflected from a concave mirror (radius of curvature 2 metres) fixed to the balance beam. The field of the permanent magnet, for an area of 5·6 sq. cm. round and at right angles to the axis, had an average value of 232 gauss per sq. cm.

The apparatus as supplied by the makers was modified in a number of details.

Screw arrangements were introduced for the adjustment of the damping magnet and for the adjustment of the hook supporting the torsion wire. The torsion wire used was about 33 cm. long, and was at first of platinum, of diameter 0·15 mm., 0·10 mm., or 0·075 mm., according to requirements, and was later of quartz, of diameter 0·10 mm. The substitution of quartz greatly improved the steadiness of the measurements. The source of light used was a small G.E.C. Osram automobile lamp (2 volts, 1·50 watts). The filament of this lamp is supplied in two forms: a spiral of five turns, or a single wave. The five-fold spiral is the more suitable. A special lamp-holder and light-projector were used to suit this lamp.

The deflexion was measured by the movement of a spot of light on a circular scale.

A strip of translucent paper, 6 cm. broad and 2 metres long, was bent into the form of a vertical circular cylinder of radius equal to about 2 metres. On this strip of paper was a scale 1·5 cm. broad and 2 metres long, divided into mm. This long scale was very useful in the adjusting of the apparatus before taking a series of measurements, and also in the preliminary approximate measurements with substances for which the susceptibility was of an entirely unknown order of magnitude.

Only a small portion in the centre of this long scale was required for the determination of the susceptibility—from a few mm. to about 20 cm.

Instead of the arrangement with two reins for controlling the circular movement of the permanent magnet, an arrangement with one rein was substituted, whereby a micrometric control was obtained.

The whole balance, including the case, together with the cylinder placed on the roof of the case and supporting the torsion wire, was put into a small enclosure. The incident and reflected beams of light passed through a small opening in the wall of the enclosure, a small door being opened during the taking of a measurement. The temperature of the enclosed space was kept more or less constant by means of an arrangement of electrical heaters. The average temperature was about 17° C. This method of controlling the temperature of the whole balance has a great advantage over any method whereby the control is applied only to that portion of the balance where the substance is situated. Further, the substance to be measured can be allowed to sit in the enclosed space long enough to ensure that it has had time to acquire the necessary temperature before the measurement is taken.

The measurements were determined relative to distilled water, the susceptibility of which was taken as $-0{\cdot}72 \times 10^{-6}$.

All the solids were examined in powder form.

The small tube supplied by the makers for holding the substance is of glass of very small susceptibility. This property is of the greatest value for this work. We found that this glass was, in this respect, much superior to quartz.

The deflexion produced by the empty tube was checked before each filling, and a deflexion for water was obtained in most cases before each measurement of the deflexion for the substance under examination. This was necessary because unaccountable changes in the deflexion for water were obtained from time to time.

The Preparation and Purification of the Substances.

In order to enhance the value of the magnetic measurements an effort was made to obtain the highest purity. The presence of a very small amount of impurity (especially if paramagnetic) in a diamagnetic substance may, of course, lead to quite erroneous results. Chemical analysis, although useful as a method of identification, is of less service as a criterion of purity than the melting-point, which is again surpassed for this purpose by the measurement of the susceptibility itself. In many cases no useful object is served by carrying out a chemical analysis. The best

procedure in such cases is to adopt some method or methods of purification and to measure any improvement in purity by means of the melting-point, or boiling-point, or density and the susceptibility. An improvement in purity not discernible by analysis or melting-point may be disclosed by a change in susceptibility, the values for which not only alter but also become more regular as the purity grows better.

Phenyl Halides.

Phenyl iodide.—A specimen obtained from British Drug Houses was redistilled several times under reduced pressure and also at atmospheric pressure. The middle portion of the distillate was collected (b.-p. 186–186·5° C.). Redistillation did not alter the susceptibility to any great extent. The molecular susceptibility obtained from an average of four measurements was considerably below Pascal's experimental value ($-90{\cdot}16 \pm 0{\cdot}51$ against Pascal's $-96{\cdot}0$).

Phenyl chloride.—The liquid was carefully redistilled. Four measurements gave a molecular susceptibility equal to $-69{\cdot}61 \pm 0{\cdot}35$, or 1·6 unit below Pascal's experimental value.

Iodobenzene dichloride.—Dry chlorine was passed through a solution of iodobenzene in chloroform (5 grams in 10 c.c. chloroform). Crystals of $C_6H_5ICl_2$ separated. These were filtered off and washed with chloroform. The compound was recrystallized from chloroform. The first measurements of susceptibility were not satisfactory, being too irregular. A second preparation gave better agreement among the several values: mol. sus. $= -117{\cdot}6 \pm 1{\cdot}2$ (5).

Halides of Rubidium.

Rubidium iodide.—This was obtained from British Drug Houses, and examined magnetically without further treatment. Mol. sus. $= -67{\cdot}05 \pm 1{\cdot}9$ (4).

RbI_3—This was prepared by the method of Wells and Wheeler (*Zeit. f. Anorg. Chem.* i. p. 443, 1892).

2 grams RbI and 2·2 grams iodine were dissolved by warming in about 2 c.c. water. The solution was cooled in ice; when almost black glistening crystals separated. These were filtered off, washed with a little alcohol, and dried. When dry, the crystals were powdered in an agate mortar. A red-brown powder was obtained. Magnetic measurements

gave unsatisfactory results. A paramagnetic impurity was suspected. Analysis, by decomposing with heat, showed the presence of traces of an impurity insoluble in water.

RbBr.—The material obtained from British Drug Houses was measured magnetically without being recrystallized. Mol. sus. was found to be $-56{\cdot}69 \pm 0{\cdot}54$ (4).

The sum of Pascal's atomic values gives $-57{\cdot}6$.

The experimental value of N. Crow (Trans. Roy. Soc. Canada, xix. 1925) is $-62{\cdot}9$.

$RbIBr_2$.—This was prepared by the method of Wells and Wheeler (*Zeit. f. Anorg. Chem.* i. p. 443, 1892).

To a saturated solution of 2 grams RbBr in water was added 1·5 gram iodine and 1 gram bromine. The mixture was warmed till all went into solution. The solution was cooled in ice, when crystals of $RbIBr_2$ separated. These were filtered off, washed with alcohol, and dried.

A weighed quantity of the substance was heated in a crucible till all the added bromine and iodine was removed. The residue was then weighed.

Loss in weight = 54·2 per cent. (average of two determinations).

Theoretical value = 55·57 per cent.

Mol. sus. $= -116{\cdot}2 \pm 0{\cdot}53$ (2).

Halides of Tetramethylammonium.

The polyhalides of this series was prepared by the methods of Chattaway and Hoyle (Journal of Chemical Society, cxxiii. pp. 654–662, 1923).

$N(CH_3)_4I$.—Material from Kahlbaum was recrystallized from water and washed with alcohol and dried.

Magnetic measurements before and after recrystallization agreed pretty closely with each other. Mol. sus. $= 105 \pm 0{\cdot}88$ (3).

$N(CH_3)_4I_3$.—Equivalent amounts of iodine and $N(CH_3)_4I$ were dissolved in boiling alcohol. Crystals of $N(CH_3)_4I_3$ separated on cooling along with some unchanged $N(CH_3)_4I$. The $N(CH_3)_4I_3$ was extracted with cold alcohol, and the solution was then evaporated to saturation. The crystals obtained were washed with alcohol and dried. The m.-p. given by Chattaway was obtained: 118° C. Three magnetic measurements were made with this product, which was then recrystallized, and a further three measurements were carried out. After recrystallization there was an increase of about

0·01 in the specific susceptibility and 4·6 in the molecular susceptibility. The last three values were regarded as more accurate, and their average was taken as the result. Mol. sus. $= -181{\cdot}9 \pm 1{\cdot}6$ (3).

$N(CH_3)_4I_5$. — Equivalent amounts of $N(CH_3)_4I_3$ and iodine were dissolved in hot alcohol. The solution was allowed to cool when crystals of $N(CH_3)_4I_5$ separated. These were washed with alcohol and dried. M.-p. was 130° C., somewhat lower than Chattaway's 132° C.

Only enough material was available for one magnetic measurement, and this, together with the low melting-point, leads us to place little confidence in the magnetic result, which does not square with theory.

$N(CH_3)_4Br$.—Material from Kahlbaum was recrystallized several times throughout the measurements from a mixture of alcohol and water. The crystals were washed with alcohol, dried at the pump and in a vacuum desiccator. The magnetic figures were raised by recrystallization, but the several values did not agree well among themselves. Measurements with a fresh quantity of the substance gave concordant susceptibilities which agreed also with several of the values got for the first sample. These values were used to give the mol. sus. $-87{\cdot}21 \pm 0{\cdot}46$ (6).

$N(CH_3)_4I_2Br$.—Equivalent amounts of $N(CH_3)_4Br$ and iodine were dissolved in boiling alcohol. The required compound separated on cooling. It was recrystallized from alcohol. We did not succeed in reaching Chattaway's m.-p. 136° C. Our first m.-p. was 126° C., and repeated recrystallization both from alcohol and from acetic acid gave finally m.-p. 127° C. A second preparation of the trihalide gave m.-p. 126° C., and after recrystallization m.-p.'s of 129° to 130° C. This compound was therefore of doubtful purity.

The specimen (m.-p. 129° C.) gave specific susceptibilities $-0{\cdot}4115$ and $0{\cdot}4108$, and that with m.-p. 130° C. gave $-0{\cdot}3966$, $-0{\cdot}3921$, and $-0{\cdot}3973$. The average of the latter three values was taken as the best available result. Mol. sus. $= -161{\cdot}3 \pm 0{\cdot}89$ (3).

$N(CH_3)_4IBr_2$.—1 c.c. of bromine dissolved in 2 c.c. of glacial acetic acid was added to 3·6 grams of $N(CH_3)_4I$ suspended in 45 c.c. of alcohol. The mixture was heated till solution was effected. The required crystals separated on cooling. They were recrystallized from alcohol. A second crop of crystals was recrystallized from acetic acid.

The magnetic results were very concordant. The crystals from acetic acid gave approximately the same value as those from alcohol. Mol. sus. $= -150{\cdot}8 \pm 0{\cdot}78$ (4).

$N(CH_3)_4Br_3$.—A quantity of $N(CH_3)_4Br$ was dissolved in the smallest possible quantity of a hot mixture of equal volumes of alcohol and acetic acid. Bromine in small excess was then added and the solution allowed to cool. The crystals obtained were recrystallized from alcohol containing a little bromine. This compound is not very stable. The m.-p. was 117–118° C. (Chattaway 118·5°C.). The magnetic results obtained before and after recrystallization were of the same order, but not very concordant. Mol. sus. $= -137{\cdot}0 \pm 2{\cdot}1$ (4).

$N(CH_3)_4BrCl_2$.—Dry chlorine was passed through a solution of $N(CH_3)_4Br$ in the smallest quantity of acetic acid, contained in a conical flask surrounded by a freezing mixture. Crystals separated and were filtered off and dried at the pump. Only one magnetic measurement was made. Mol. sus. $= -114{\cdot}1$ (1).

Halides of Tetraethylammonium.

The compounds examined in this series were prepared by the methods of Chattaway and Hoyle (Journal of Chemical Society, cxxiii. pp. 654–662, 1923).

$N(C_2H_5)_4I$.—Material from Kahlbaum was recrystallized from alcohol. Magnetic measurements were carried out before and after recrystallization. The average of the latter values were taken for the result. Mol. sus. $= -152{\cdot}6 \pm 0{\cdot}31$ (3).

$N(C_2H_5)_4I_3$.—Equivalent quantities of $N(C_2H_5)_4I$ and iodine were dissolved in hot alcohol. The product is not very soluble in alcohol, and separated readily as the solution cooled. The m.-p. was 142° C.

A considerable number of magnetic determinations were carried out with the products of several recrystallizations in order to obtain concordant results. Good agreement, however, was not got, and the average of five values of specific susceptibility ranging from $-0{\cdot}433$ to $-0{\cdot}449$ was taken. Mol. sus. $= -224{\cdot}9 \pm 2{\cdot}9$ (5).

$N(C_2H_5)_4IBr_2$.—To a quantity of $N(C_2H_5)_4I$ suspended in alcohol an equivalent quantity of bromine was added. The mixture was warmed till all went into solution. It was then allowed to cool, when crystals of $N(C_2H_5)_4IBr_2$ separated.

The m.-p. was 124° C. (Chattaway 125° C.). A large number of magnetic measurements were made, and recrystallization produced no change in the values obtained. The average of nine values was taken. Mol. sus. $= -202{\cdot}7 \pm 0{\cdot}87$ (9).

The Halides of Phenyltrimethylammonium.

All the compounds examined in this series, except $C_6H_5N(CH_3)_3I_3$, were obtained from preparations by Mr. T. H. Reade (McCombie and Reade, Journal of Chemical Society, cxxiii. pp. 141–153 (1923).

$PhNMe_3I$.—This substance was recrystallized from alcohol. Two magnetic measurements were carried out before recrystallization and two after. The latter two agreed well with the former two. Compare

		Diamagnetic gain.
	$NH_3 + HI \rightarrow NH_4I$	3·14
and		
	$NPhMe_2 + MeI \rightarrow NPhMe_3I$.....	0·6

Mol. sus. of $PhNMe_3I = -144 \pm 3{\cdot}2$ (4).

$PhNMe_3I_3$—Equimolecular quantities of $PhNMe_3I$ and iodine were dissolved in the necessary volume of acetic acid by heating. On cooling, crystals of $PhNMe_3I_3$ separated. These were filtered off and recrystallized from acetic acid. The m.-p. was 115–116° C. McCombie and Reade gave 116°.

Three magnetic measurements were made on the product of this preparation. After further recrystallization two more measurements were made, and after a subsequent recrystallization a sixth determination was carried out. The recrystallization produced only very small changes in the susceptibility. The average of six values was taken. Mol. sus. $= -211{\cdot}2 \pm 5{\cdot}9$ (6).

The magnetic values for the other compounds of this series that were examined are given in Table I.

The Chlorides of Iodine.

The magnetic measurements given in Table I. and referred to elsewhere in the paper, were obtained in this laboratory by Dr. John Farquharson, who prepared and analysed the substances measured. On passing chlorine over iodine, at first ICl is obtained in the form of a reddish-brown liquid, and subsequently, *i. e.* with excess of chlorine, ICl_3 appears as yellow crystals. According to requirements, the reaction

could be stopped at the ICl stage, with the use of a special apparatus, or allowed to go to the end, when very simple apparatus sufficed. The progress of the reaction had to be tested by analysis, and the purity of the materials that were measured could be checked either by analysis or by the degree of steadiness in the measurements or by both these means.

The following results were obtained :—

ICl.—

Percentage of chlorine.	Mol. sus.
Theory 21·84; found 22·24	54·61
22·49	54·18
22·01	54·84
Average...	54·54 ± 0·24 (3)

ICl_3.—When chlorine had been passed for three hours after the yellow crystal stage was reached, analysis gave 44·84 per cent. of Cl. Theory requires 45·61. It was necessary to use freshly sublimed material for each measurement in order to keep the purity up to that standard required for concordant resultants. Material kept for a few days showed a lowered susceptibility. The result from the best material was :—Mol. sus. $= 90{\cdot}16 \pm 0{\cdot}69$ (4).

J. D. desires to thank The Department of Scientific and Industrial Research for allowance during one year of the period occupied in this research.

Chemistry Department,
Aberdeen University.

V. *An Effect of Space Charge in a Gas at Low Pressures.* *By* E. W. B. Gill, *B.Sc., M.A., Fellow of Merton College, Oxford**.

WITH an ordinary three-electrode valve as used for wireless purposes it has been found under certain conditions that, when definite potentials are applied to the electrodes, the current passing through the valve from filament to anode can have two or three different values †. This effect is more readily obtained if the applied potentials are small (from 10 to 30 volts).

* Communicated by Prof. J. S. Townsend, F.R.S.
† E. W. B. Gill, Phil. Mag. xlix. p. 993 (May 1925).

The potential distribution in the space between the electrodes is different with the different currents, although the potentials of the electrodes themselves are the same.

A like effect can be obtained with a two-electrode valve if it contains gas at low pressures, the applied potential between the anode and the filament being again small and of the order of that usually quoted as the ionization potential of the gas.

The two-electrode valve with which the following results were obtained consisted of a filament surrounded by a coaxial cylinder of approximately circular section, the mean diameter being about 2·5 cm. The valve contained helium at a pressure of 8 mm. of mercury. It was connected with a tube containing charcoal cooled with liquid air to remove impurities from the gas. The emission currents from the filament to the positively charged cylinder for different potentials between the electrodes were measured while the filament was kept at a bright red heat.

As usual, the currents remained very small till the potential difference between the electrodes reached about 20 volts, after which a rapid rise in the current occurred. If the changes in potential are made by breaking the circuit and changing the number of accumulators in the battery the currents obtained with increasing potentials are exactly the same as those obtained with decreasing potentials; but if the potentials are varied continuously by a potentiometer this is no longer the case. The currents in the table are an example of this, the potential having been increased continuously from 22 to 28 volts, and then reduced to 22 volts.

PO in volts	22	23	24	25	26	27	28
Current in MA increasing.....	·08	·12	·38	·96	1·74	2·36	4·9
Current in MA decreasing....	2·86	3·42	3·72	3·74	3·8	3·8	

It will be noticed that as the potential is reduced from 28 volts there is a reduction of the current from 4·9 to 3·8 milliamps at 27 volts, and there is no further large reduction in current until the potential is reduced to 22 volts. When the potential is reduced below 22 volts the current drops suddenly to the original small value.

If at any stage in the reduction of potential the circuit is momentarily broken and remade the current returns to the ascending voltage value. In order to get different currents with any potentials between 22 and 28 volts by this method it is necessary to bring the potential up to 28 volts and then reduce it.

It is, however, possible to obtain the higher currents for the lower voltages without going first up to 28 volts. It has been found that a magnetic field applied parallel to the filament while the smaller current is running causes the current to increase immediately to a higher value, and when the magnetic field is removed the current remains at the high value given in the third line of figures in the table.

This is of importance as indicating that the effect is in no way dependent on the higher potentials being first applied, which might put the gas in some abnormal condition, and the following simple general explanation seems adequate.

In all cases the currents passing are limited by the negative space charges, and the increases of current when ionization occurs are due not so much to the actual magnification of the current by ionization, as to the neutralization by the positive ions of part of the negative charge in the gas.

The effect of the negative space charge is to reduce the electric force near the filament and increase it near the anode. Ionization will only occur when the electrons have reached a certain limiting velocity, and the above distribution of electric force reduces the space over which the electrons are capable of ionizing atoms of the gas. If by any chance the large current shown in line 3 is set up this indicates that the number of positive ions is large, the negative space charge is largely neutralized, and the region over which the electrons can ionize is increased. The necessary larger supply of positive ions will thus be maintained.

In any case, the possibility of two currents for any potential difference in the region of the ionizing potentials in the simple case of only two electrodes indicates the difficulty of interpreting results obtained with more complicated arrangements.

If the temperature of the filament be reduced so as to make the emission of the electrons very small the effect

of the space charge becomes inappreciable, and there is only one anode current for a given potential difference between the electrodes.

It will be noticed that, the pressure being 8 mm., each electron will make several thousand collisions with atoms of the gas in passing from the filament to the anode. The total amount of energy thus lost is of the order of 8 or 10 volts, but nevertheless quite a large proportion attain the energy required to ionize when the total potential is only 22.

VI. *On the Rolling of a Heavy Solid on a Horizontal Plane in the Absence of both Sliding and Spinning.* *By* M. A. Omara, *Docteur ès Sciences (Grenoble), Lecturer in Applied Mathematics in the Egyptian University* *.

§ 1. *Introduction.*

THE problem of the rolling of a heavy body on a horizontal plane was considered by Slesser ('Quarterly Journal of Mathematics,' 1861) and by Routh ('Advanced Rigid Dynamics'), who, independently of one another, wrote the equations of motion. Application to the particular problem of the rolling of a disk has been given by Ferrers ('Quarterly Journal of Mathematics,' 1872), by Carvallo (*J. de l'Ecole Polytechnique*, 2e série, 1900–1901), and by Kortweg, and by Appell (*Rendiconti del Circolo math. di Palermo*, 1899). Also Appell, in the *J. de Math.* 1900, dealt with the rolling of a solid of revolution on a horizontal plane. But in all these works spinning was considered to take place at the same time as the rolling, *i.e.*, the rolling friction and the spinning friction were both neglected with respect to the sliding friction.

The present writer, in a paper published in the *Annales de l'Université de Grenoble*, 1925, compared these two kinds of friction, and pointed out circumstances in which the rolling friction can be neglected with respect to that of spinning, and then studied the rolling without spinning of a solid of revolution on a horizontal plane.

In this paper the general case of rolling without spinning for a heavy body is considered, and the equations of motion are reduced to two by the use of Appell's "energy of acceleration" function. These equations are then integrated

* Communicated by Prof. A. M. Mosharrafa.

in the case of a solid of revolution. The general form of the curve described by the point of contact on the surface is studied, and application to the case of an ellipsoid of revolution is given.

§ 2. *Kinematical Considerations.*

Consider a heavy solid bounded by a convex surface Σ, and moving, under gravity, in contact with a horizontal plane H. Let $Gxyz$ be a moving system formed by the principal axes of inertia at the centre of gravity, and $O\xi\eta\zeta$ a fixed system. $O\zeta$ is the upward vertical and $O\xi$, $O\eta$ two rectangular axes in the plane H. The position of the solid is defined by five parameters, namely, the three Eulerian angles θ, ϕ, ψ, and ξ, η the coordinates of the projection of the centre of gravity on the plane H. The ordinate ζ is a function $F(\theta, \phi)$ which can be easily formed when the equation of the surface Σ is given. Let this equation be $g(xyz)=0$. Let x, y, z be the coordinates of the point of contact P of the solid with the plane H, and u, v, w the components of the velocity of the centre of gravity, and p, q, r the components of the angular velocity of the solid. Let also α, β, γ be the direction-cosines of the upward vertical at the point P. α, β, γ are functions of x, y, z, given by the well-known relations:

$$\frac{\alpha}{\frac{\partial g}{\partial x}} = \frac{\beta}{\frac{\partial g}{\partial y}} = \frac{\gamma}{\frac{\partial g}{\partial z}} = \frac{-1}{\sqrt{\left(\frac{\partial g}{\partial x}\right)^2 + \left(\frac{\partial g}{\partial y}\right)^2 + \left(\frac{\partial g}{\partial z}\right)^2}}, \qquad \text{(N)}$$

since the direction of the upward vertical is that of the inward normal to the surface at the point of contact.

The geometrical conditions of the constraint—*i.e.*, the contact of the solid with the horizontal plane—are expressed by writing

$$\left.\begin{aligned} \dot{\alpha} + \gamma q - \beta r &= 0, \\ \dot{\beta} + \alpha r - \gamma p &= 0, \\ \dot{\gamma} + \beta p - \alpha q &= 0, \end{aligned}\right\} \quad \ldots\ldots \quad (1)$$

and

$$\zeta + \alpha x + \beta y + \gamma z = 0, \quad \ldots\ldots \quad (2)$$

and the kinematical conditions expressing the absence of sliding and spinning are respectively

$$\left.\begin{aligned} u + qz - ry &= 0, \\ v + rx - pz &= 0, \\ w + py - qx &= 0, \end{aligned}\right\} \quad \ldots\ldots \quad (3)$$

and

$$p\alpha + q\beta + r\gamma = 0. \quad \ldots\ldots \quad (4)$$

If $G(\alpha, \beta, \gamma, \zeta) = 0$ be the tangential equation of the surface Σ, the coordinates x, y, z of the point P are expressed n terms of α, β, γ by the equations

$$x \cdot \frac{\partial G}{\partial \zeta} = \frac{\partial G}{\partial \alpha}; \quad y \cdot \frac{\partial G}{\partial \zeta} = \frac{\partial G}{\partial \beta}; \quad z \cdot \frac{\partial G}{\partial \zeta} = \frac{\partial G}{\partial \gamma}$$

Also the components p, q, r of the angular velocity can be expressed as functions of the parameters α, β, γ and their derivatives in a symmetrical manner: multiply the second of equations (1) by γ and the third by $-\beta$ and add, we find

$$\gamma\dot{\beta} - \beta\dot{\gamma} = p(\beta^2 + \gamma^2) - \alpha(\beta q + \gamma r),$$

but
$$\alpha^2 + \beta^2 + \gamma^2 = 1,$$

and from equation (4)

$$\beta q + \gamma r = -\alpha p;$$

therefore

$$\left.\begin{aligned} p &= \gamma\dot{\beta} - \beta\dot{\gamma}, \\ q &= \alpha\dot{\gamma} - \gamma\dot{\alpha}, \\ r &= \beta\dot{\alpha} - \alpha\dot{\beta}, \end{aligned}\right\} \quad \ldots\ldots \quad (5)$$

also

and
$$p^2 + q^2 + r^2 = \alpha^2 + \dot{\beta}^2 + \dot{\gamma}^2.$$

Differentiating (5) with respect to time, we have

$$\left.\begin{aligned} \dot{p} &= \gamma\ddot{\beta} - \beta\ddot{\gamma}, \\ \dot{q} &= \alpha\ddot{\gamma} - \gamma\ddot{\alpha}, \\ \dot{r} &= \beta\ddot{\alpha} - \alpha\ddot{\beta}. \end{aligned}\right\} \quad \ldots\ldots \quad (5\,a)$$

Multiplying these equations respectively by α, β, γ, and adding, we find

$$\dot{p}\alpha + \dot{q}\beta + \dot{r}\gamma = 0. \quad \ldots\ldots \quad (5\,b)$$

Also, substituting for p, q, r from equations (5) in (3), we find for the first

$$u = \dot{\alpha}(y\beta + z\gamma) - \alpha(y\dot{\beta} + z\dot{\gamma}),$$

and on differentiating equation (2) with respect to time, we find

$$\dot{\zeta} + \dot{\alpha}x + \dot{\beta}y + \dot{\gamma}z = 0,$$

because
$$\alpha\dot{x} + \beta\dot{y} + \gamma\dot{z} = 0,$$

as we can see from the relations (N). Therefore we have

$$u = -\dot{\alpha}(\zeta + \alpha x) + \alpha(\dot{\zeta} + x\dot{\alpha}),$$

i.e.,
$$u = \alpha\dot{\xi} - \zeta\dot{\alpha},$$
also
$$v = \beta\dot{\xi} - \zeta\dot{\beta},$$
$$w = \gamma\dot{\xi} - \zeta\dot{\gamma},$$
. . . . . (6)

and

$$u^2 + v^2 + w^2 = \dot{\xi}^2 + \zeta^2(\dot{\alpha}^2 + \dot{\beta}^2 + \dot{\gamma}^2)$$
$$= \dot{\xi}^2 + \zeta^2(p^2 + q^2 + r^2). \quad . \quad . \quad (6\,a$$

Differentiating equations (6) with respect to time, we find

$$\dot{u} = \alpha\ddot{\xi} - \zeta\ddot{\alpha},$$
$$\dot{v} = \beta\ddot{\xi} - \zeta\ddot{\beta},$$
$$\dot{w} = \gamma\ddot{\xi} - \zeta\ddot{\gamma}.$$
. . . . . (6 *b*)

§ 3. *Equations of Motion.*

Taking the same moving axes G*xyz*, as in § 2, let A, B, C be the principal moments of inertia. Now, equations (3) and (4) being non-integrable, the system is non-holonomic. We shall use Appell's general equations of motion (*loc. cit.*). Let

$$2S = \Sigma m_i(\ddot{x}_i^2 + \ddot{y}_i^2 + \ddot{z}_i^2),$$

where m_i typifies the mass of a particle of the moving body, whose rectangular coordinates are x_i, y_i, z_i. This function S is called by French writers the "énergie d'accélération" of the system. It is a quadratic function of the quantities $\ddot{q}_1, \ddot{q}_2, \ldots \ddot{q}_k$, where $\ddot{q}_n (n = 1, 2, \ldots k)$ are the independent parameters, k being the degree of freedom of the system. Any small displacement of the system which is consistent with the constraints can be defined by changes $(\delta q_1, \delta q_2, \ldots \delta q_k)$ of the independent parameters. The work done by the external forces in such a displacement is

$$\delta T = \Sigma(X\delta x + Y\delta y + Z\delta z)$$
$$= \sum_1^k P_n\, \delta q_n.$$

Appell's equations of motion are then

$$\frac{\partial S}{\partial \ddot{q}_n} = P_n \qquad (n = 1, 2, \ldots k).$$

The function S admits of a theorem analogous to that of König with regard to the kinetic energy of a system. This theorem is

$$S = \tfrac{1}{2}MJ_0^2 + S_1,$$

where J_0 is the acceleration of the centre of gravity and S_1 the energy of acceleration in the motion relative to the centre of gravity, and S the total energy of acceleration.

The projections of the acceleration of the centre of gravity on the moving axes are

$$J_{0,x} = \dot{u} + qw - rv, \quad J_{0,y} = \ldots, \quad J_{0,z} = \ldots,$$

and

$$\begin{aligned} J_0^2 &= J_{0,x}^2 + J_{0,y}^2 + J_{0,z}^2 \\ &= \dot{u}^2 + \dot{v}^2 + \dot{w}^2 + 2p(v\dot{w} - w\dot{v}) + 2q(w\dot{u} - u\dot{w}) \\ &\qquad + 2r(u\dot{v} - v\dot{u}) + \ldots . \end{aligned}$$

Terms not involving derivatives of the second order of the parameters are omitted for brevity. Also, substituting for $\dot{u}$, $\dot{v}$, and $\dot{w}$ from (6 *b*), and noting that

$$\alpha\dot{\alpha} + \beta\dot{\beta} + \gamma\dot{\gamma} = 0,$$

and that

$$\begin{aligned} \alpha\ddot{\alpha} + \beta\ddot{\beta} + \gamma\ddot{\gamma} &= -(\dot{\alpha}^2 + \beta^2 + \dot{\gamma}^2) \\ &= -(p^2 + q^2 + r^2), \end{aligned}$$

we may write

$$\begin{aligned} J_0^2 = \ddot{\zeta}^2 + \zeta^2(\ddot{\alpha}^2 + \ddot{\beta}^2 + \ddot{\gamma}^2) + 2\zeta\{\ddot{\alpha}(rv - qw) + \ddot{\beta}(pw - ru) \\ + \ddot{\gamma}(qu - pv)\} + \ldots \end{aligned}$$

In order to calculate S_1, let m be a particle whose co-ordinates are x, y, z with respect to the system $Gxyz$. The components of its absolute velocity are

$$v_x = qz - ry\ ; \quad v_y = \ldots\ ; \quad v_z = \ldots .$$

The components of its acceleration are

$$J_x = \dot{v}_x + qv_z - rv_y\ ; \quad J_y = \ldots\ ; \quad J_z \ldots .$$

Substituting for v_x, v_y, v_z, we find

$$J_x = -x(q^2 + r^2) + z\dot{q} - y\dot{r} + p(qy + rz),$$

and similar expressions for J_y and J_z. Therefore

$$\begin{aligned} 2S_1 &= \Sigma m(J_x^2 + J_y^2 + J_z^2) \\ &= A\dot{p}^2 + B\dot{q}^2 + C\dot{r}^2 + 2(C - B)qr\dot{p} + 2(A - C)rp\dot{q} \\ &\qquad + 2(B - A)pq\dot{r} + \ldots , \end{aligned}$$

omitting terms not containing derivatives of the second order. Putting the mass of the body $M = 1$, we have

$$2S = \ddot{\zeta}^2 + \zeta^2(\ddot{\alpha}^2 + \ddot{\beta}^2 + \ddot{\gamma}^2) + 2\zeta[\ddot{\alpha}(rv - qw) + \ddot{\beta}(pw - ru) + \ddot{\gamma}(qu - pv)] + A\dot{p}^2 + B\dot{q}^2 + C\dot{r}^2 + 2(C - B)qr\dot{p} + 2(A - C)rp\dot{q} + 2(B - A)pq\dot{r} + \dots .$$

If, now, we take as independent variables the Eulerian angles θ and ϕ, we have

$$\alpha = \sin\theta\sin\phi, \quad \beta = \sin\theta\cos\phi, \quad \gamma = \cos\theta,$$

and

$$\zeta = F(\theta, \phi),$$

so that equations (5), (5 *a*), (6), and (6 *b*) give p, q, r, u, v, w and their derivatives in terms of θ and ϕ and their derivatives of the first and second orders. The function S is thus expressed as a function of $\ddot{\theta}$ and $\ddot{\phi}$.

Now, in a small displacement consistent with the constraints existing at time t, the independent variables undergo the increments $\delta\theta$ and $\delta\phi$. The work done by the external forces during this displacement reduces to the work done by the weight of the body, *i. e.*,

$$\delta T = -g\delta\zeta = -g\left[\frac{\partial F}{\partial\phi}\delta\phi + \frac{\partial F}{\partial\theta}\delta\theta\right].$$

The equations of motion are

$$\left.\begin{aligned}\frac{\partial S}{\partial\ddot{\phi}} &= -g\frac{\partial F}{\partial\phi},\\ \frac{\partial S}{\partial\ddot{\theta}} &= -g\frac{\partial F}{\partial\theta}.\end{aligned}\right\} \quad \dots\dots \quad (7)$$

When integrated these equations give θ and ϕ as functions of the time. Equation (4), which may be written in terms of the Eulerian angles θ, ϕ, and ψ,

$$p\alpha + q\beta + r\gamma = \dot{\psi} + \dot{\phi}\cos\theta = 0,$$

will determine ψ.

Equations (7) admit of the energy integral, which may be written

$$u^2 + v^2 + w^2 + Ap^2 + Bq^2 + Cr^2 = h - 2g\zeta.$$

Let us write, *in extenso*, the first of equations (7); we have

$$\frac{\partial S}{\partial\ddot{\phi}} = \ddot{\zeta}\frac{\partial\ddot{\zeta}}{\partial\ddot{\phi}} + \zeta^2\left(\ddot{\alpha}\frac{\partial\ddot{\alpha}}{\partial\ddot{\phi}} + \ddot{\beta}\frac{\partial\ddot{\beta}}{\partial\ddot{\phi}} + \ddot{\gamma}\frac{\partial\ddot{\gamma}}{\partial\ddot{\phi}}\right) + \zeta\left[(rv - qw)\frac{\partial\ddot{\alpha}}{\partial\ddot{\phi}} + (pw - ru)\frac{\partial\ddot{\beta}}{\partial\ddot{\phi}} + (qu - pv)\frac{\partial\ddot{\gamma}}{\partial\ddot{\phi}}\right]$$

$$+[A\dot{p}+(C-B)qr]\frac{\partial\dot{p}}{\partial\ddot{\phi}}+[B\dot{q}+(A-C)rp]\frac{\partial\dot{q}}{\partial\ddot{\phi}}$$

$$+[C\dot{r}+(B-A)pq]\frac{\partial\dot{r}}{\partial\ddot{\phi}}=-g\frac{\partial F}{\partial\phi}.$$

But we notice that

$$\frac{\partial\ddot{\alpha}}{\partial\ddot{\phi}}=\beta,\quad \frac{\partial\ddot{\beta}}{\partial\ddot{\phi}}=-\alpha,\quad \frac{\partial\ddot{\gamma}}{\partial\ddot{\phi}}=0,$$

and that

$$\frac{\partial\ddot{\xi}}{\partial\ddot{\phi}}=-x\frac{\partial\ddot{\alpha}}{\partial\phi}-y\frac{\partial\ddot{\beta}}{\partial\phi}-z\frac{\partial\ddot{\gamma}}{\partial\phi}=y\alpha-x\beta;$$

also

$$\frac{\partial\dot{p}}{\partial\phi}=\gamma\frac{\partial\ddot{\beta}}{\partial\ddot{\phi}}-\beta\frac{\partial\ddot{\gamma}}{\partial\ddot{\phi}}=-\gamma\alpha$$

$$\frac{\partial\dot{q}}{\partial\phi}=-\gamma\beta,\ \frac{\partial\dot{r}}{\partial\phi}=\alpha^2+\beta^2=\lambda^2\ (\text{say});$$

so that the first of equations (7) reduces to

$$\ddot{\zeta}(y\alpha-x\beta)+\zeta^2(\ddot{\alpha}\beta-\ddot{\beta}\alpha)+\zeta\{r\,v\beta+u\alpha)-w(p\alpha+q\beta)\}$$

$$-[A\dot{p}+(C-B)qr]\gamma\alpha-[B\dot{q}+(A-C)rp]\gamma\beta$$

$$+[C\dot{r}+(B-A)pq]\lambda^2=-g\frac{\partial F}{\partial\phi}.$$

But, from equation (4),

$$p\alpha+q\beta=-r\gamma,$$

and from equations (6)

$$\alpha u+\beta v+\gamma w=\zeta.$$

Hence, we finally obtain

$$(y\alpha-x\beta)+\zeta^2\dot{r}+r\zeta\dot{\zeta}-[A\dot{p}+(C-B)qr]\gamma\alpha$$

$$-[B\dot{q}+(A-C)rp]\gamma\beta+[C\dot{r}+(B-A)pq]\lambda^2=-g\frac{\partial F}{\partial\phi}.$$

. . . . (8)

In the same manner the second of equations (7) may be written

$$\ddot{\xi}\{z\sin\theta-\cos\theta(x\sin\phi+y\cos\phi)\}$$

$$+\zeta^2\{\ddot{\theta}-\dot{\phi}^2\sin\theta\cos\theta\}+\zeta\dot{\xi}\dot{\phi}+[A\dot{p}+(C-B)qr]\cos\phi$$

$$-[B\dot{q}+(A-C)rp]\sin\phi=-g\frac{\partial F}{\partial\theta}. \quad . \quad . \quad . \quad . \quad (9)$$

In equations (8) and (9) ζ, p, q, r and their derivatives and x, y, z are to be replaced by their values in terms of θ and ϕ and their derivatives of the first and second orders. These values are easily computed, as the quantities in question are already given as functions of α, β, γ, which are in their turn known functions of θ and ϕ.

§ 4. *Solution for a Homogeneous Solid of Revolution.*

In the case of a homogeneous solid of revolution take the axis of revolution as Gz; Gx and Gy will be any two perpendicular axes in the equatorial plane. Let

$$g(\sqrt{x^2+y^2}, z) = 0$$

be the equation of the surface of revolution bounding the solid; the tangential equation of this surface will be

$$\psi(\sqrt{\alpha^2+\beta^2}, \gamma, \zeta) = 0,$$

or, what is the same thing,

$$\psi(\lambda, \gamma, \zeta) = 0.$$

Solving for ζ, we get

$$\zeta = f(\lambda, \gamma) = f(\sin\theta, \cos\theta).$$

We see that the height of the centre of gravity above the horizontal plane is independent of the angle ϕ; so that

$$\frac{\partial F}{\partial \phi} = \frac{\partial f}{\partial \phi} = 0.$$

In order to calculate the coordinates x, y, z of the point of contact P in terms of the angles θ and ϕ, write the tangential equation of the surface in the form

$$f(\lambda, \gamma) - \zeta = G(\alpha, \beta, \gamma, \zeta) = 0.$$

Then

$$\frac{\partial G}{\partial \zeta} = -1, \quad \frac{\partial G}{\partial \alpha} = \frac{\partial f}{\partial \lambda}\cdot\frac{\partial \lambda}{\partial \alpha} = \frac{\partial}{\partial \lambda}\cdot\frac{\alpha}{\lambda},$$

$$\frac{\partial G}{\partial \beta} = \frac{\partial f}{\partial \lambda}\cdot\frac{\partial \lambda}{\partial \beta} = \frac{\partial f}{\partial \lambda}\cdot\frac{\beta}{\lambda}, \quad \frac{\partial G}{\partial \gamma} = \frac{\partial f}{\partial \gamma};$$

and therefore

$$x = -\frac{\alpha}{\lambda}\frac{\partial f}{\partial \lambda}, \quad y = -\frac{\beta}{\lambda}\frac{\partial f}{\partial \lambda}, \quad z = -\frac{\partial f}{\partial \gamma}.$$

We see hat, in the case of a solid of revolution,

$$y\alpha - x\beta = 0.$$

Equation (8) is now written

$$\zeta^2\dot{r}+r\dot{\zeta}\zeta-A\gamma\{\dot{p}\alpha+\dot{q}\beta\}+(A-C)r\gamma(q\alpha-p\beta)+C\lambda^2\dot{r}=0.$$

But, from equation (5 *b*), $\dot{p}\alpha+\dot{q}\beta=-\dot{r}\gamma$, and from the third of equations (1) $q\alpha-p\beta=\dot{\gamma}$. Also $\lambda^2+\gamma^2=1$, and therefore $\gamma\dot{\gamma}=-\lambda\dot{\lambda}$.

Equation (8) finally becomes

$$\dot{r}[A\gamma^2+C\lambda^2+\zeta^2]+r[A\gamma\dot{\gamma}+C\lambda\dot{\lambda}+\zeta\dot{\zeta}]=0.$$

Integrating, we have

$$r^2[A\gamma^2+C\lambda^2+\zeta^2]=r_0^2[A\gamma_0^2+C\lambda_0^2+\zeta_0^2]=R_0^2. \quad (10)$$

We have thus arrived at a first integral which will permit the reduction of the problem to quadratures.

The energy integral can be written, in virtue of equation (6 *a*),

$$\dot{\zeta}^2+(A+\zeta^2)(p^2+q^2)+(C+\zeta^2)r^2=h-2g\zeta.$$

But

$$p^2+q^2=\frac{1}{\lambda^2}(\gamma^2r^2+\dot{\gamma}^2),$$

therefore

$$\dot{\zeta}^2+\frac{1}{\lambda^2}[(A+\zeta^2)\dot{\gamma}^2+R_0^2]=h-2g\zeta,$$

i.e., $$\left[A+\zeta^2+\left(\frac{\partial\zeta}{\partial\theta}\right)^2\right]\sin^2\theta\,.\,\dot{\theta}^2=(h-2g\zeta)\sin^2\theta-R_0^2. \quad (11)$$

If $\zeta=f(\gamma,\lambda)$ be a rational function of λ and γ, this equation will give t as a function of θ by a hyperelliptic quadrature, when we take $\tan\frac{\theta}{2}$ as the independent variable.

The third of equations (5) gives

$$r=\beta\dot{\alpha}-\alpha\dot{\beta}=\dot{\phi}\,.\,\lambda^2.$$

Substituting for r from equation (10), we have

$$\dot{\phi}=\frac{R_0}{\lambda^2\sqrt{A\gamma^2+C\lambda^2+\zeta^2}}. \quad . \quad . \quad . \quad . \quad (12)$$

Eliminating dt between this equation and equation (11), we obtain ϕ as a function of θ also by a quadrature.

As

$$\dot{\psi}+\dot{\phi}\cos\theta=\dot{\psi}+\dot{\phi}\gamma=0,$$

therefore

$$\dot{\psi}=\frac{-\gamma R_0}{\lambda^2\sqrt{A\gamma^2+C\lambda^2+\zeta^2}}. \quad . \quad . \quad . \quad . \quad (13)$$

Eliminating dt between this equation and (11), we get ψ as a function of θ by another quadrature.

§ 5. *Form of the Curve described by the Point of Contact on the Solid of Revolution.*

Suppose that the body is bounded by a convex surface formed by the revolution of a continuous convex curve round the z-axis. This curve is assumed to possess no singular points nor points at infinity. ζ, the height of the centre of gravity above the tangent plane, as well as $\frac{\partial \zeta}{\partial \theta}$, which represents the distance of the centre of gravity from the normal to the surface at the point of contact, will both be finite.

Now, the left-hand side of equation (11) is essentially positive. In order that $\dot{\theta}$ may be real the right-hand side must also be positive during the motion. Put

$$\delta(\theta) = (h - 2g\zeta)\sin^2\theta - R_0^2.$$

The function $\delta(\theta)$ is negative for $\theta = 0$ and $\theta = \pi$. It is necessarily positive for the initial value θ_0 of θ, which lies between 0 and π. We see that the equation $\delta(\theta) = 0$ has two real roots, the one θ_1 lies between 0 and θ_0, and the other between θ_0 and π. The angle θ librates between θ_1 and θ_2 throughout the motion. When it reaches the one or the other of these two values $\dot{\theta}$ vanishes and changes its sign.

The locus of the point of contact P lies between two parallels whose colatitudes are θ_1 and θ_2. When the initial conditions are such that the roots θ_1 and θ_2 coincide they have the common value θ_0. $\dot{\theta}$ will then be always naught, and the point of contact describes a parallel of colatitude θ_0.

From equation (12) we see that $\dot{\phi}$ is never zero, *i.e.*, it keeps a constant sign, and the solid rotates always in the same sense round its axis of revolution.

Now call V the angle that the tangent to the path of the point P on the surface makes with the meridian passing through P. To a small displacement PP′ of the point of contact corresponds a small arc $PP_1 = \rho d\phi$ of the parallel passing through P and an arc $PP_2 = ds$ of the meridian. We have

$$\tan V = \frac{\rho d\phi}{ds} = \rho \frac{d\phi}{dt}\frac{dt}{ds},$$

ρ being the radius of the parallel at the point of contact, which is always finite and different from zero, as θ is never zero nor π. The same is true for $\frac{d\phi}{dt}$. The quantity $\frac{ds}{dt}$ which represents the velocity of the point of contact P in the direction of the tangent to the meridian passing through P is necessarily finite. We see that $\tan V$ never vanishes; also the locus of P presents neither cusps nor loops.

We may write

$$\frac{ds}{dt} = \frac{ds}{d\theta}\frac{d\theta}{dt};$$

the factor

$$\frac{ds}{d\theta}\sqrt{\left(\frac{d\rho}{d\theta}\right)^2 + \left(\frac{dz}{d\theta}\right)^2}$$

is never naught, as the meridian presents no cusp. $\frac{ds}{dt}$ vanishes at the same time as $\frac{d\theta}{dt}$, *i. e.*, only when θ reaches one of the roots θ_1 or θ_2; $\tan V$ then becomes infinite and the locus of P is tangent to the parallels of colatitudes θ_1 and θ_2.

§ 6. *Case of an Ellipsoid of Revolution.*

Consider a heavy ellipsoid of revolution whose centre of gravity is at its geometric centre. We have

$$\zeta^2 = a^2\sin^2\theta + c^2\cos^2\theta,$$

where a and c are the semi-axes of the meridian.

Substituting for $\sin\theta$, $\cos\theta$, and $\dot\theta$ in terms of ζ and $\dot\zeta$ in equation (11), it becomes

$$(\mathfrak{A}\zeta^2 - a^2c^2)\dot\zeta^2 = (a^2-\zeta^2)[(h-2g\zeta)(\zeta^2-c^2) - R_0^2(a^2-c^2)],$$

where $\mathfrak{A} = A + a^2 + c^2$.

t is given as a function of ζ by a hyperelliptic quadrature. An elementary study of the roots of the polynomial

$$\delta(\zeta) = (h-2g\zeta)(\zeta^2-c^2) - R_0^2(a^2-c^2)$$

shows that $\delta(\zeta) = 0$ has three roots—one, ζ_1, essentially negative, the other two, ζ_2 and ζ_3, both positive. ζ_2 lies between zero and ζ_0.

If $c > a$, *i. e.*, if the axis of revolution is the major axis, ζ must remain in the interval (a, ζ_2), where $\zeta_2 < c$. The angle θ oscillates between θ_2 and $(\pi-\theta_2)$, θ_2 being the value of θ corresponding to ζ_2; and thus the locus of the point of contact has a sinusoidal form (fig. 1).

The two roots ζ_2 and ζ_3 may coincide and have the common value c. The angle θ will then oscillate between zero and π. But the solid will take an infinite time to attain the one or the other of these two positions. The locus of P will have

Fig. 1.

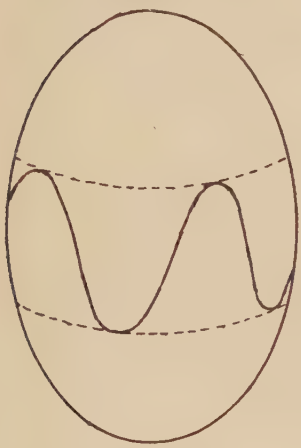
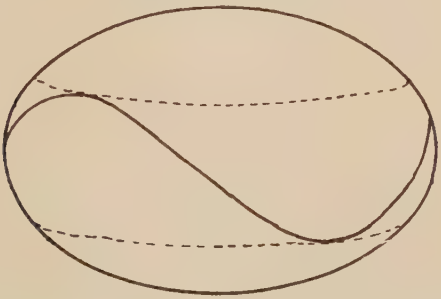


Fig. 2.

the form of a double spiral surrounding the ellipsoid. The asymptotical points will be those where the axis of revolution meets the ellipsoid (fig. 2).

If $a > c$, *i. e.*, if the ellipsoid is of revolution round the minor axis, θ may oscillate between the two values θ_2 and

Fig. 3.



$(\pi-\theta_2)$, as in the preceding case. The locus of P will have the same form (fig. 3). But it may also happen that the third root, ζ_3, is smaller than a. ζ will then lie between the two values ζ_2 and ζ_3, and the angle θ will oscillate either in the interval (θ_2, θ_3) or in the interval $\{(\pi-\theta_3), (\pi-\theta_2)\}$.

The locus of P will then be composed of two sinusoidal parts, symmetrical with respect to the equator (fig. 4).

If $\zeta_3 = a$, the position defined by $\theta = \frac{\pi}{2}$ is an asymptotical one, and an infinite time is spent before it is arrived at. The locus of P will again have the form of a double spiral, with the circle of the equator as an asymptotical circle (fig. 5).

Fig. 4.

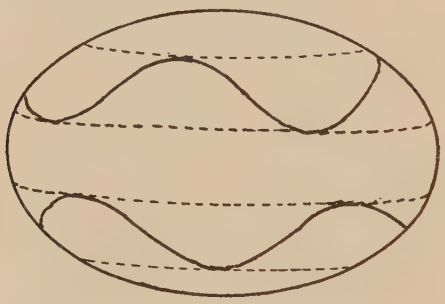
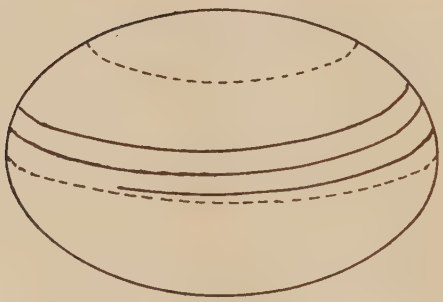


Fig. 5.



Summary.

The motion of a heavy solid on a horizontal plane is discussed, and dynamical equations are derived, using Appell's "energy of acceleration" function. These equations are reduced to quadratures in the case of a solid of revolution. The nature of the path of the point of contact is investigated, and application to the special case of an ellipsoid of revolution is given.

The author's thanks are due to Prof. A. M. Mosharrafa for his kindly help and criticism.

VII. *Properties of Saturated Vapours, including Steam.* *By* Prof. A. Press *.

Summary.

It had been thought that no equation of state could be derived from the two fundamental laws of thermodynamics. The present work shows in what manner the first law (or energy equation) can be solved (in an improved manner), giving rise to the two functional equations

$$pv = (\Gamma - 1)(E - V),$$

$$\left(p + \frac{1}{v^{\Gamma}}\int v^{\Gamma-1}\,\frac{dV}{dv}\,dv\right)v = \frac{\Gamma-1}{v^{\Gamma-1}}\psi(tv^{\Gamma-1}).$$

The first includes the "fundamental equation" used by Callendar for the construction of his steam tables, as a special case ; the second leads to a theoretical justification of Nernst's "Third Law of Thermodynamics," among other things.

It is shown that for saturated vapours the equation of state necessarily takes the form

$$p = (\Gamma - 1)\,.\,a_m\,.\,t^{\frac{\Gamma}{\Gamma-1}},$$

with the internal energy given by the expression

$$E = \beta + a_m\,.\,v\,.\,t^{\frac{\Gamma}{\Gamma-1}},$$

where β is Callendar's constant for the mixture region. Making use therefore of the latter determinations, it is shown that the ratio of the gross latent heat of evaporation L to the net internal latent heat of evaporation $\mathscr{L}$ is given by

$$L/\mathscr{L} = \Gamma = (\text{constant characterizing the fluid}).$$

Further checking the above two formulæ for saturated steam by means of steam tables, with vapour pressures ranging from 4 mm. to 17,387 mm. of mercury, temperatures 0° C. to 220° C., the constancy of Γ for steam is proved to be practically,

$$\Gamma = 1{\cdot}11.$$

The latter agrees rather well with Zeuner's estimate 1·135 and John Perry's estimate, based on experiments with steam engines, 1·14. The two formulæ can thus be employed quite generally in getting out the "properties of saturated vapours."

* Communicated by the Author.

§ 1. *The Universal Characteristic and General Equation of State.*

THE energy equation thermodynamically being

$$dQ = dE + pdv, \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

it is a simple matter to show, if μ is an integrating factor, then, according to the criterion condition of Euler for a perfect differential, we must have necessarily

$$\frac{\partial}{\partial v}\Big[\mu\Big]_E = \frac{\partial}{\partial E}\Big[\mu p\Big]_v \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

The latter equation can be solved in a relatively simple way by making use of the transformation

$$F = lg(\mu W), \text{ where } W = W(v) \text{ only.} \quad . \quad . \quad (3)$$

It can be shown, for example, that the solution of (3) comprises the two equations functionally

$$pv = (\Gamma - 1)(E - V), \quad . \quad . \quad . \quad . \quad . \quad (4)$$

$$\left(p + \frac{1}{v^\Gamma}\int v^{\Gamma-1}.dV\right)v = \frac{\Gamma-1}{v^{\Gamma-1}}.\overline{\psi}\left(\frac{\mu}{v^{\Gamma-1}}\right). \quad . \quad . \quad (5)$$

The first, (4), represents the universal characteristic, and includes Callender's so-called "Fundamental Equation" as a special case applying to the *mixture region* with

$$V = V(v) = \beta. \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

On the other hand, the second, or (5), gives a theoretical basis to all equations of state of the van der Waals type. The two expressions (4) and (5) are improved forms resulting from a simplified procedure that is given herewith.

Operating with (3), we have that

$$\frac{\left(\frac{\partial F}{\partial v}\right)_E}{\left(\frac{\partial F}{\partial E}\right)_v} = \frac{\frac{1}{\mu}\left(\frac{\partial \mu}{\partial v}\right)_E + \frac{1}{W}.\left(\frac{\partial W}{\partial v}\right)_E}{\frac{1}{\mu}\left(\frac{\partial \mu}{\partial E}\right)_v} \quad . \quad . \quad . \quad . \quad (7)$$

Again, equation (2) can quite easily be thrown into the form

$$p = \frac{\frac{1}{\mu}\left(\frac{\partial \mu}{\partial v}\right)_E - \left(\frac{\partial p}{\partial E}\right)_v}{\frac{1}{\mu}\left(\frac{\partial \mu}{\partial E}\right)_v} \quad . \quad . \quad . \quad . \quad . \quad (8)$$

Correlating (7 and (8), it follows, setting

$$\left(\frac{\partial p}{\partial \mathrm{E}}\right)_v = -\frac{1}{\mathrm{W}}\left(\frac{\partial \mathrm{W}}{\partial v}\right)_{\mathrm{E}}, \quad . \quad . \quad . \quad . \quad (9)$$

that the problem of solving (2) becomes dependent on being able to evaluate the equation

$$\left(\frac{\partial \mathrm{F}}{\partial v}\right)_{\mathrm{E}} = p\left(\frac{\partial \mathrm{F}}{\partial \mathrm{E}}\right)_v. \quad . \quad . \quad . \quad . \quad . \quad (10)$$

In other words, equation (2) is thereby broken up into two subsidiary equations (9) and (10).

The convenience and importance of this procedure is that, solving for (9), we have

$$\left.\begin{aligned} p &= \left\{ -\frac{\partial}{\partial v}(lg\,\mathrm{W})_{\mathrm{E}} \right\}_{11}(\mathrm{E}-\mathrm{V}), \\ \mathrm{V} &= \mathrm{V}(v) \text{ only}, \end{aligned}\right\}, \quad . \quad . \quad (11)$$

where E is the internal energy as before and V, as the "constant of integration," takes on the nature of a cohesional internal energy term which in the hands of Callendar was taken to be

$$\mathrm{V} \equiv \beta \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (12)$$

for the steam mixture region. In fact (11) really is Callendar's so-called "Fundamental Equation" generalized. From a *dimensional* point of view we need to have

$$p = \{\}_{11}^{-1} \equiv \text{energy}, \quad . \quad . \quad . \quad . \quad . \quad (13)$$

just as E and V must represent energy. In other words, *the simplest development* * to take of $\{\}_{11}$ to satisfy the dimensions of energy is

$$\left.\begin{aligned} -\frac{\partial}{\partial v}(lg\,\mathrm{W}) &= \frac{1}{v} \\ \text{with } \mathrm{W} &= \frac{1}{v^{\Gamma-1}} \end{aligned}\right\}. \quad . \quad . \quad . \quad (14)$$

The universal characteristic following from the energy equation (1) thus becomes

$$pv = (\Gamma-1)(\mathrm{E}-\mathrm{V}). \quad . \quad . \quad . \quad . \quad (15)$$

* This does not mean *dimensionally* that more complex developments may not at times be necessary.

Callendar, in his "Steam Tables," adopts the notation

$$p(v-b) = \frac{1}{n}(u-\beta). \quad . \quad . \quad . \quad . \quad . \quad (16)$$

(See, for example, Roberts, 'Heat and Thermodynamics,' p. 436.) There is left, however, equation (10) to be disposed of.

It will be found that writing (10) in the form

$$\left(\frac{\partial F}{\partial v}\right)_E = (\Gamma-1).\frac{E-V}{v}.\left(\frac{\partial F}{\partial E}\right)_v, \quad . \quad . \quad (17)$$

because of (15) just obtained, that the solution of (17) is given as

$$F = F\{(E-V)v^{\Gamma-1} + \int v^{\Gamma-1}.dV\}. \quad . \quad . \quad (18)$$

The method employed in arriving at this result is a "divisional" one, already disclosed by the writer in the Trans. Roy. Soc. Canada, sect. iii. p. 365 (1927). If, then, the functional solution (18) is interpreted backwards in the light of (3) with respect to (14), the general equation of state becomes

$$pv = \frac{\Gamma-1}{v^{\Gamma-1}}\left[\overline{\psi}\left(\frac{\mu}{v^{\Gamma-1}}\right) - \int v^{\Gamma-1}.dV\right]. \quad . \quad . \quad (19)$$

With the classical integrating factor taken as

$$\mu = \frac{1}{t} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (20)$$

it means

$$\frac{pv}{\Gamma-1} = E-V = \Sigma a_m.t^m.v^{(m-1)(\Gamma-1)} - \frac{1}{v^{\Gamma-1}}.\int v^{\Gamma-1}.dV. \quad (21)$$

For condensed systems, where Nernst's Third Theorem had its original successes, it would be necessary to interpret Σ and $\overline{\psi}$ as follows :—

$$E-V = \sum_0^m a_m t^m.v^{(m-1)(\Gamma-1)} - \frac{1}{v^{\Gamma-1}}.\int v^{\Gamma-1}.dV \quad . \quad . \quad (22)$$

For the vapour (and gaseous) state, however, we rather have that

$$E-V = \sum_1^m a_m t^m.v^{(m-1)(\Gamma-1)} - \frac{1}{v^{\Gamma-1}}.\int v^{\Gamma-1}.dV. \quad . \quad . \quad (23)$$

§ 2. *The Ratio of Gross to Net Internal Latent Heats of Evaporation.*

Firstly, with regard to the theoretical justification of Callendar's formula. Since the cohesional function V is

really dependent on the density of the phase where $\rho = 1/v$, it should naturally follow that for the mixture region the V value should be constant for the vapour as well as the liquid. The universal characteristic thus becomes

$$pv = (\Gamma - 1)(E - \beta), \quad . \quad . \quad . \quad . \quad . \quad (24)$$

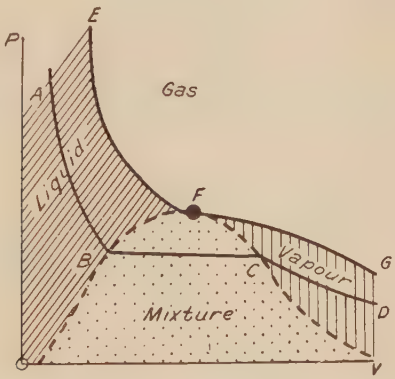
and, moreover, in view of (22) and (23) we have that the integral vanishes, for then

$$\frac{1}{v^{\Gamma-1}} \cdot \int v^{\Gamma-1} . dV \equiv 0. \quad . \quad . \quad . \quad . \quad (25)$$

Equation (24) will therefore hold for steam.

To arrive at the equation of state for the mixture region it needs to follow that we have

$$p = F(t), \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (26)$$



where t is the absolute temperature. An inspection of the figure shows why this should be so for the (p, v)-plane. Making use, therefore, of solution (21), it means we have to choose the m's in such a manner because of (26) that the resulting equation shall give

$$(\Gamma - 1)\Sigma a_m t^m . \frac{v^{m(\Gamma-1)}}{v^{\Gamma}} \equiv F(t) \equiv p. \quad . \quad . \quad (27)$$

To satisfy a relationship such as (27) it follows by the method of indeterminate coefficients

$$\left.\begin{aligned} m(\Gamma - 1) &= \Gamma, \\ m &= \frac{\Gamma}{\Gamma - 1}, \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad (28)$$

in order to avoid any v-terms. The equation of state and the internal energy functions become

$$\left.\begin{aligned} p &= (\Gamma-1).a_m.t^{\frac{\Gamma}{\Gamma-1}}*, \\ \mathrm{E} &= \beta + a_m.v.t^{\frac{\Gamma}{\Gamma-1}}. \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad (29)$$

These equations are to hold for the mixture region.

The first application of equations (29) will be toward finding the ratio of the gross to the net latent heat of evaporation. The definition of the net latent heat of evaporation is given by the relation

$$\mathscr{L} = \int_v \lambda.dv = \int_{\substack{v_l = \text{liquid.} \\ v_v = \text{vapour.}}} \left(\frac{\partial \mathrm{E}}{\partial v}\right)_t .dv. \quad . \quad . \quad . \quad (30)$$

Substituting (29) in (30) we have that

$$\mathscr{L} = a_m.t^{\frac{\Gamma}{\Gamma-1}}.(v_v - v_l), \quad . \quad . \quad . \quad . \quad (31)$$

where t is the temperature at which the evaporation takes place. On the other hand, the gross-value L depends on finding

$$\begin{aligned} \mathrm{L} = \int_v \left\{ p + \left(\frac{\partial \mathrm{E}}{\partial v}\right)_t \right\} dv &= \int_v pdv + \mathscr{L} \\ &= \int_v pdv + \int_v \left(\frac{\partial \mathrm{E}}{dv}\right)_t .dv. \quad . \quad (32) \end{aligned}$$

We have, however, that

$$\int_v pdv = (\Gamma-1).a_m.t^{\frac{\Gamma}{\Gamma-1}}.(v_v - v_l). \quad . \quad . \quad (33)$$

The ratio of the gross to the net is therefore determined by the formula

$$\mathrm{L}/\mathscr{L} = 1 + (\Gamma-1) = \Gamma. \quad . \quad . \quad . \quad . \quad (34)$$

The Edser observation of constancy of $\mathrm{L}/\mathscr{L}$ (see Fourth Report on Colloid Chemistry, Brit. Assoc. Adv. Sci. 1922) thus becomes well founded. It will moreover be noted that Zeuner, in his experimental work on steam engines, showed that the equation of expansion was given by

$$\left.\begin{aligned} pv^{\Gamma} &= \text{constant} \\ \text{where } \Gamma &\equiv 1{\cdot}135. \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad (35)$$

* As a matter of record Bertrand's suggestion, quoted by Partington, 'Chemical Thermodynamics,' p. 60, gives $p = \mathrm{A}\{(t-\beta)/t\}^n$. It is known to give good results. See also equation (10 *b*) of Roberts (*l. c.*), p. 178.

The derivation of Γ as the adiabatic index was given in my previous paper (Phil. Mag. (Dec. 1924), or see *Zeit. f. Phys.* p. 138 (1929). John Perry, in his 'Steam Engine,' p. 362, gives rather

$$\Gamma \equiv 1{\cdot}14. \quad . \quad . \quad . \quad . \quad . \quad . \quad (36)$$

§ 3. *Formula for the Properties of Saturated Vapours and Steam.*

By adding the result (33) above to the evaluation $\mathcal{L}$ in (31), it follows for the gross latent heat of evaporation that

$$\mathrm{L} = \frac{a_m}{\Gamma} \cdot (v_v - v_l) \cdot t^{\frac{\Gamma}{\Gamma-1}}. \quad . \quad . \quad . \quad . \quad (37)$$

At the critical point (see F of the figure) the vapour volume v_v becomes equal to the liquid volume v_l, so that the latent heat value L becomes zero for this critical condition. Nevertheless, in checking formula (37) by means of published steam tables, the liquid volume v_l will be ignored in comparison with the all-vapour volume v_v. The constant a_m can of course be determined by taking actual values of L, v_v, v_l, and t once the Γ-value has been found. For this reason a tabulation is appended based on the data for L, v_v, and t only. The type of calculation necessary is as follows :—

Taking two sets of values obtaining for formula (37), it should follow on division that

$$\frac{\mathrm{L}}{\mathrm{L}'} \equiv \frac{v_v}{v_v'} \cdot \left(\frac{t}{t'}\right)^{\frac{\Gamma}{\Gamma-1}}. \quad . \quad . \quad . \quad . \quad . \quad (38)$$

Thus, operating with log on both sides we have

$$\log\left(\frac{\mathrm{L}}{\mathrm{L}'} \cdot \frac{v_v'}{v_v}\right) = \frac{\Gamma}{\Gamma-1} \cdot \log\left(\frac{t}{t'}\right). \quad . \quad . \quad . \quad (39)$$

Writing for convenience

$$\chi = \frac{\log\left(\frac{\mathrm{L}}{\mathrm{L}'} \cdot \frac{v_v'}{v_v}\right)}{\log\left(\frac{t}{t'}\right)}, \quad . \quad . \quad . \quad . \quad . \quad (40)$$

it is easily proved

$$\frac{\Gamma}{\Gamma-1} = \chi; \quad \Gamma = \frac{\chi}{\chi-1}. \quad . \quad . \quad . \quad . \quad (41)$$

The calculations were made according to formula (40) from a vapour pressure of practically only 4 mm. of mercury to a

vapour pressure 17,387 mm. of mercury. The temperature range was actually from 0° C. up to 220° C. without making any allowance for the liquid volume originally obtaining before evaporation began. The last column in (42) shows to what extent a formula such as (37) can be employed in projecting steam tables or the like.

Saturated Steam.

°C.	$t°$.	L= calories per kilo.	v= cubic metres per kilo.	p=mm. Hg.	Γ= calculated base 100° C.
0	273	595·4	206·3	4·6	1·07
100	373	538·7	1·671	760	
150	423	504·1	·3921	3569	1·1
200	473	463·8	·1274	11,647	1·11
220	493	446	·0860	17,387	1·11

Jan. 1930.

VIII. *Classical and Modern Gravitational Theories.* *By* Prof. A. Press *.

Summary.

In a previous paper ("Classical and Modern Electromagnetic Theories," Phil. Mag. Nov. 1929), to which the present acts as a sequel, it was shown definitely that the major problems which gave rise to the Special Theory of Relativity could be explained classically without having to resort to the non-Newtonian space-time of Einstein. Moreover, the outstanding Miller-Residual was explained classically—a point which the Relativity Theory with its Lorentzian Transformation could not explain.

In the present paper the so-called "crucial problems" are taken up, and again it is shown that there is no real necessity of introducing either the relativistic space-time or the Lorentzian transformation. As a matter of fact, in the earlier paper (*l. c.* pp. 642–643) a very important error was shown to have crept into the Lorentzian analysis, due to

* Communicated by the Author.

disregarding the essential limitation under which div ρu could be equated to $d\rho/dt$. As a result the invariant transformations made so much of in General Relativity Theory fall by the wayside. Thus it becomes even more important to establish a classical basis for Einstein's "Theory of Gravitation," implying thereby his orbital equation

$$\frac{d^2u}{d\theta^2}+u=\frac{m}{h^2}+\frac{2}{3}ku^2,$$

together with the Bending of Light and Spectral Shift enunciations.

It is shown that the two classical acceleration equations,

$$\frac{d^2r}{dt^2}-r\left(\frac{d\theta}{dt}\right)^2=-\frac{\mathrm{P}}{m};\quad \frac{1}{r}\frac{d}{dt}\left(r^2\frac{d\theta}{dt}\right)=-\mathrm{Q},$$

lead to Einstein's equation under the supposition that, instead of the simple Keplerian law, we assume that

$$r^2\frac{d\theta}{dt}=h-\frac{1}{2}\cdot\frac{q.h^3}{c^2r^3}.$$

It is proved that the additional term $\frac{1}{2}qh^3/c^2r^3$ implies that in a gravitational field a solid body is given a rotation or gravitational spin, as it were, which should lengthen and shorten the day as a function of the time-rate of change of the radius vector. The deduced relation is that

$$\mathrm{Q}=\frac{3}{2}\cdot\frac{qh^3}{c^2r^5}\cdot\frac{dr}{dt}.$$

There is a decided advantage in the present analysis as against that used by Einstein, inasmuch as the coefficient q does not have to be inflexibly taken equal to k, so that there is room for adjustment with regard to the observed perihelion of Mercury motion. Of course, where an orbit is substantially circular, as is the case with Venus, no variation from the Newtonian law should occur. A further advantage is that the Spectral Shift evaluation becomes very much smaller than that indicated by Einstein's General Relativity Theory.

The remaining problem, that of the Bending of Light, has been treated by Einstein on the assumption of a "Principle of Equivalence." This non-classical principle has been avoided. In its place the assumption is put forth that possibly the trajectory of a ray of light is affected in some ("corpuscular," Newton) manner by the æther strain set up by the gravitational influence of the sun's mass.

Indeed, an orbital type of equation is developed, based on using Gaussian coordinates in a classical manner—that is, without necessarily introducing a Lorentz-Einstein invariance. In fact the time t is not made an independent variable after the non-classical relativistic manner. Thus, for a purely Newtonian gravitational field the Gaussian equation subsisting is shown to be

$$dr^2+r^2\left(1-\frac{2mr}{h^2}\right)d\theta^2=c^2dt^2,$$

the subsidiary equation being

$$r^2\left(\frac{d\theta}{dt}\right)=h.$$

In the previous equation c is the velocity of light in the free æther. Making $m=0$ then leads to the expression

$$d\sigma^2=c^2dt^2.$$

The latter is a Newtonian result, and indicates that a length of arc $d\sigma$ can be measured by the employment of a light signal, a time-measuring clock, and a photographic camera. In the Theory of General Relativity this type of experiment is not to be conceived, since the Einsteinian Invariant ds^2 has to be reckoned with. The Einsteinian "Interval" ds, with its accompanying independent variable t, are *per contra* excluded from the present consideration. Nevertheless, a Duality Principle does appear to be involved. Thus, taking instead of the Keplerian law,

$$r^2\frac{d\theta}{dt}=h,$$

the more general value involving q, then Einstein's gravitational equation follows from the classical Gaussian equation instead of the Plym equation previously given. The "duality principle" is, then, nothing more than assuming that in a gravitational field a light ray is refracted so as to make it act like a Newtonian "corpuscle." That is, we need to take $m=0$ in the Gaussian equation classically derived or alternatively in Einstein's gravitational equation!

It is then by means of the classical Gaussian equation and the generalized Keplerian law that the Einsteinian "Bending of Light" can be shown to follow. However, for the Spectral Shift the law deduced is rather

$$\frac{\mathrm{T\,(earth)}}{\mathrm{T\,(sun)}}=1+\frac{1}{2}\frac{qh^2}{c^2}\left(\frac{1}{\mathrm{R}_e^3}-\frac{1}{\mathrm{R}_s^3}\right),$$

whereas Einstein's corresponding formula is

$$\frac{\mathrm{T\,(earth)}}{\mathrm{T\,(sun)}} = 1 + \frac{1}{2}k\left(\frac{1}{\mathrm{R}_e} - \frac{1}{\mathrm{R}_s}\right).$$

A spectral shift effect is actually indicated, but its magnitude is very much less than that following from Relativistic Principles. The above are the results of a classical interpretation of Einstein's philosophic contributions.

---

§ 1. THE General Theory of Relativity is in particular characterized by application to three crucial problems so-called. The one has to do with the Perihelion of Mercury, the second the Bending of Light, and the third the Spectral Shift. These three problems will be investigated from a classical point of view, and it will be shown, as in the previous paper ("Classical and Modern Electromagnetic Theories," Phil. Mag., Nov. 1929), that there is no need for departing from the Newtonian ideas of time and space to arrive substantially at Einstein's results.

Conceive, therefore, of the following preliminary problem, in which it is sought to measure a length r by means of a time-measuring clock together with a light impulse or ray sent from one end to the other, assuming as known the velocity of light c. If instead of r the length dr is employed, it will be proper to express the law of action by writing classically that

$$\left.\begin{aligned} dr^2 &= c^2\,dt^2, \\ c &= \frac{dr}{dt}. \end{aligned}\right\} \quad \ldots\ldots \quad (1)$$

In this case, therefore, it would be quite proper to regard $\frac{dr}{dt}$ in (1) as measuring the rate with which a light ray would travel along the length from an origin where the light pulse was being sent out to a further point r, where a camera and appropriate clock were placed to record photographically the arrival time t of the pulse. (A second order correction would be necessary, as pointed out in the previous paper (*l. c.*), if the physical length to be measured moved relative to the wave-sustaining medium.) Instead of the equations (1) of course it would be equally proper, adopting a length of arc $d\sigma$, to write instead of dr

$$d\sigma^2 = dr^2 + r^2\,d\theta^2 = c^2.dt^2. \quad \ldots\ldots \quad (2)$$

Thus, by division it would then normally follow,

$$\left(\frac{dr}{d\theta}\right)^2 + r^2 = c^2\left(\frac{dt}{d\theta}\right)^2 . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

The Gaussian equation (2) thus particularizes the length-measuring problem in which a light ray is being employed.

§ 2. Consider now a second type of problem, namely, the orbital one with central forces in which Newtonian gravitation applies. It is well known that the motion is symbolized by the two equations

$$\frac{d^2u}{d\theta^2} + u = \frac{m}{h^2}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

$$r^2\frac{d\theta}{dt} = h. \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

It is desired to express equation (4) in the form of a Gaussian relationship between the variables which shall at one and the same time include the possibility described in § 1. (This is, in essence, the striving after an analogous "Unit Field Theory.") Take, therefore, the differential characteristic

$$dr^2 + r^2\left(1 - \frac{2mr}{h^2}\right)d\theta^2 = c^2dt^2. \quad . \quad . \quad . \quad (6)$$

Dividing through with $d\theta^2$ we have that

$$\left(\frac{dr}{d\theta}\right)^2 + r^2 - \frac{2mr^3}{h^2} = c^2\left(\frac{dt}{d\theta}\right)^2 . \quad . \quad . \quad . \quad (7)$$

If, then, we operate on $u = \frac{1}{r}$, substitutionally, it follows

$$\frac{dr}{d\theta} = -\frac{1}{u^2}\cdot\frac{du}{d\theta}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

$$\frac{d\theta}{dt} = \frac{h}{r^2} = hu^2. \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

Introducing (8) and (9) into (6) or (7) gives at once that

$$\frac{1}{u^4}\cdot\left(\frac{du}{d\theta}\right)^2 + \frac{1}{u^2} - \frac{2m}{h^2}\cdot\frac{1}{u^3} = \frac{c^2}{h^2u^4}, \quad . \quad . \quad (10)$$

$$\left(\frac{du}{d\theta}\right)^2 + u^2 - \frac{2m}{h^2}\cdot u = \frac{c^2}{h^2}. \quad . \quad . \quad . \quad (11)$$

On differentiating (11) with respect to θ the right-hand

side, involving the light velocity c, disappears, which is significant, and we have

$$2\frac{du}{d\theta}\cdot\frac{d^2u}{d\theta^2}+2u.\frac{du}{d\theta}=\frac{2m}{h^2}\cdot\frac{du}{d\theta}. \quad . \quad . \quad . \quad (12)$$

The latter leads on simplification to the result already announced, namely

$$\frac{d^2u}{d\theta^2}+u=\frac{2m}{h^2}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

At the same time, of course, *taking m in* (6) *equal to zero*, the equation is obtained

$$dr^2+r^2.d\theta^2=c^2dt^2, \quad . \quad . \quad . \quad . \quad . \quad (2)$$

which is precisely the equation started out with in § 1. Equation (6) can therefore be interpreted in a dual manner! It is as if light classically followed corpuscularly the orbital laws (4) and (5), but with the understanding that for length-measuring purposes we need to take in conjunction with (6) that

$$m=0. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (13)$$

The duality is rather suggestive, to say the least. It enables (4) to reduce to the equation

$$\left.\begin{aligned}&\frac{d^2u}{d\theta^2}+u=0,\\ &\text{provided we take } m=0.\end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad (14)$$

There appears, then, justification for regarding this Duality Principle as analogous to a limited type of Unitary Field Theory.

§ 3. Einstein, in his gravitational theory, has not been satisfied with the Newtonian central force formulæ, (4) and (5). Instead, he has suggested an equation of the type

$$\frac{d^2u}{d\theta^2}+u=\frac{m}{h^2}+\frac{3}{2}ku^2, \quad . \quad . \quad . \quad . \quad . \quad (15)$$

which latter he employs, as is well known, to solve the perihelion of Mercury problem. The physical interpretation of such formula is that in the equations for the accelerations along a radius vector and at right angles thereto we need to set

$$\left.\begin{aligned}&\frac{d^2r}{dt^2}-r\left(\frac{d\theta}{dt}\right)^2=-\frac{\mathrm{P}}{m},\\ &\frac{1}{r}\frac{d}{dt}\left(r^2\frac{d\theta}{dt}\right)=-\mathrm{Q}.\end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (16)$$

In the classical equations (16) P would be the gravitational force acting along the radius vector, and mQ would correspond to an Einsteinian force at right angles thereto, which the sequel will show implies that a slight turning moment is given to a body by the gravitational field depending on the factor $\frac{dr}{dt}$. In other words, when the orbit is circular, as with the planet Venus, there can be no gravitational spin induced, and therefore no Einsteinian correction will take place. The detailed correlation of (16) with (15) will be taken up in a later section. At the same time it is this gravitationally induced spin that CLASSICALLY is the feature (or result) of the Einstein analysis. It needs to be noted, moreover, that the time t in (16) is not an independent variable, though strangely enough it is treated as such in Einstein's General Relativity Theory. If, then, an equation such as (6) can lead to the proper result in length measuring theory, Einstein maintains in effect by the DUALITY PRINCIPLE that equation (6) can again be employed even when Q enters into the equations (16). The matter will be further discussed below.

§ 4. Integrating the Q-equation of (16), we have that

$$h-\int Qr.dt=h' \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (17)$$

$$r^2\frac{d\theta}{dt}=h' \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (18)$$

If, then, Q is not zero, it will be shown that there exists a slight departure from the Keplerian law regarding the sweeping out of equal areas in equal times by the radius vector. Adopting provisionally the evaluation

$$r^2\frac{d\theta}{dt}=h-\frac{1}{2}\frac{qh^3}{c^2r^3}=h', \quad . \quad . \quad . \quad . \quad (19)$$

the latter can be combined with the classical Gaussian equations

$$dr^2+r^2\left(1-\frac{2mr}{h^2}\right)d\theta^2=c^2dt^2, \quad . \quad . \quad . \quad . \quad (6)$$

$$\left(\frac{dr}{d\theta}\right)^2+r^2-\frac{2mr^3}{h^2}=c^2\left(\frac{dt}{d\theta}\right)^2 \quad . \quad . \quad . \quad . \quad (7)$$

Whereas then we have, as before,

$$\frac{dr}{d\theta}=-\frac{1}{u^2}\cdot\frac{du}{d\theta}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

yet, in order to transform (6) into the appropriate u-equation, it is necessary to write, in view of (19), that

$$\frac{d\theta}{dt} = \frac{h'}{r^2} = h'u^2. \quad . \quad . \quad . \quad . \quad . \quad (20)$$

Substituting (8) and (20) into (6) or (7), then, in accordance with (10) and (11),

$$\left(\frac{du}{d\theta}\right)^2 + u^2 - \frac{2mu}{h^2} = \frac{c^2}{h'^2}. \quad . \quad . \quad . \quad . \quad (21)$$

Thus, on differentiating it means that

$$2\frac{du}{d\theta}\cdot\frac{d^2u}{d\theta^2} + 2u\frac{du}{d\theta} - \frac{2m}{h^2}\cdot\frac{du}{d\theta} = -\frac{2c^2}{h'^3}\cdot\frac{dh'}{d\theta}. \quad . \quad (22)$$

Turning therefore to (19), we have

$$\frac{dh'}{d\theta} = \frac{3}{2}\frac{qh^3}{c^2r^4}\cdot\frac{dr}{d\theta} = -\frac{3}{2}\frac{qh^3}{c^2}\cdot\frac{u^4}{u^2}\cdot\frac{du}{d\theta} = -\frac{3}{2}\frac{qh^2}{c^2}\cdot u^2\cdot\frac{du}{d\theta}. \quad (23)$$

Equation (22) thus becomes

$$\frac{d^2u}{d\theta^2} + u = \frac{m}{h^2} + \frac{3}{2}\left(\frac{h}{h'}\right)^3\cdot qu^2. \quad . \quad . \quad . \quad (24)$$

Comparing the latter with Einstein's equation (15) it is seen that

$$\left(\frac{h}{h'}\right)^3 q = \left(1 - \frac{1}{2}\frac{qh^2}{c^2r^3}\right)^3 q \equiv k. \quad . \quad . \quad . \quad (25)$$

There is, then, a possibility for a flexible perihelion correction, which latter does not exist in Einstein's treatment of the subject.

The gravitational spin correction (19) leads to Einstein's orbital equation, and is yet consistent with the classical Gaussian equation

$$dr^2 + r^2\left(1 - \frac{2mr}{h^2}\right)d\theta^2 = c^2\,dt^2. \quad . \quad . \quad . \quad (6)$$

(The latter more or less plays the part of the Einsteinian equation

$$ds^2 = -\left(1 - \frac{k}{r}\right)^{-1}dr^2 - r^2\,d\theta^2 - r^2\sin^2\theta\,.\,d\phi^2 + c^2\left(1 - \frac{k}{r}\right)dt^2.)$$

§ 5. The force term mQ, it will be found, can be evaluated by means of the general equations (16). We have

$$\frac{dr}{dt} = -\frac{1}{u^2}\cdot\frac{du}{dt} = -r^2\frac{du}{d\theta}\cdot\frac{d\theta}{dt} = -h'\cdot\frac{du}{d\theta}, \quad . \quad (26)$$

$$\frac{d^2r}{dt^2} = -h'\cdot\frac{d^2u}{d\theta^2}\cdot\frac{du}{d\theta} - \frac{du}{d\theta}\cdot\frac{dh'}{dt}. \quad . \quad . \quad . \quad . \quad (27)$$

Yet from (18) it is seen that

$$\frac{dh'}{dt} = -Qr, \quad \ldots \ldots \quad (28)$$

and we have already the relationship (20). As a result of introducing (26) to (28) into the first of (16), it now follows that

$$h'^2u^2 \cdot \frac{d^2u}{d\theta^2} - \frac{Q}{u} + h'^2 \cdot u^3 = \frac{P}{m},$$

$$\frac{d^2u}{d\theta^2} + u = \frac{P}{mh'^2u^2} + \frac{Q}{h'^2u^3} \cdot \frac{du}{d\theta} \cdot \quad \ldots \quad (29)$$

Comparing the latter with equation (24), it appears at once on equating coefficients that

$$\frac{P}{mh'^2} \cdot \frac{1}{u^2} = \frac{m}{h^2}; \quad \frac{3}{2}\left(\frac{h}{h'}\right)^3 \cdot qu^2 = \frac{Q}{h'^2u^2} \cdot \frac{du}{d\theta} \cdot \quad (30)$$

On rearrangement of terms in the P-equation a possible modification of Newton's law appears with

$$\underline{P = \frac{m^2}{r^2} \cdot \left(\frac{h'}{h}\right)^2 \cdot} \quad \ldots \ldots \quad (31)$$

As for the Q-term correction to Newton's Law of Gravitation, we have by (19) that

$$h' = h - \frac{1}{2}\frac{qh^3}{c^2r^3}, \quad \ldots \ldots \quad (32)$$

which, in view of (18), leads to the equation

$$\frac{1}{2}\frac{qh^3}{c^2r^3} = \int Qr \cdot \frac{dt}{dr} \cdot dr. \quad \ldots \ldots \quad (33)$$

Differentiating both sides with respect to r gives

$$-\frac{3}{2}\frac{qh^3}{c^2r^4} = Qr\frac{dt}{dr},$$

$$\underline{Q = -\frac{3}{2}\frac{qh^3}{c^2r^5} \cdot \frac{dr}{dt} \cdot} \quad \ldots \ldots \quad (34)$$

The Einsteinian force mQ thus tends to cause a particle to swerve at right angles to the direct Newtonian line of attraction. Such force it is seen is proportional to the

velocity $\frac{dr}{dt}$ attained along the radius vector. As stated above, for purely circular orbits such as that of Venus the effect would be nil.

§ 6. Turning to the question of the "Bending of Light," by the Principle of Duality we can take corpuscularly in the classical Gaussian equation

$$dr^2+r^2\left(1-\frac{2mr}{h^2}\right)d\theta^2=c^2dt^2. \quad . \quad . \quad . \quad (6)$$

$$m=0. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (35)$$

Then on the Newtonian basis equation (6) classically becomes

$$dr^2+r^2d\theta^2=c^2dt^2. \quad . \quad . \quad . \quad . \quad . \quad (36)$$

At the same time, if the improved Einsteinian gravitational condition

$$r^2\frac{d\theta}{dt}=h'=h-\frac{1}{2}\frac{qh^3}{c^2r^3}$$

is to be imposed instead of the purely Newtonian one, the resulting orbital equation will be

$$\frac{d^2u}{d\theta^2}+u=\frac{3}{2}qu^2. \quad . \quad . \quad . \quad . \quad . \quad (37)$$

It is the latter equation that was solved by G. B. Jeffery (see Phil. Mag. Sept. 1920, or Nunn's 'Relativity and Gravitation,' p. 123) and shown to lead to the Einstein light bending prediction. The meaning of the "Principle of Duality" is, then, that with the modified law of gravitation it is ASSUMED that projected light rays act in a manner closely analogous to the projection of Newtonian massless light corpuscles even in the modified gravitational field of Einstein. Such is the classical import of Einstein's "PRINCIPLE OF EQUIVALENCE" from a refraction point of view.

§ 6. For the examination of the "Spectral Shift" problem two similar electronic orbits about an atomic nucleus may be considered situated respectively in the sun's photosphere and the earth's surface. The Einstein correction term to the Keplerian law is

$$\frac{1}{2}\frac{qh^3}{c^2}\cdot\frac{1}{r^3},$$

where we have that

$$r^2\frac{d\theta}{dt}=h\left(1-\frac{1}{2}\frac{qh^2}{c^2}\cdot\frac{1}{r^3}\right). \quad . \quad . \quad . \quad . \quad (38)$$

If, then, instead of the gravitational centre being at a distance r we have the radius R_s, the formula should be

$$r^2\left(\frac{d\theta}{dt}\right)=h\left(1-\frac{1}{2}\frac{qh^2}{c^2}\cdot\frac{1}{R_s^3}\right). \quad . \quad . \quad . \quad . \quad (39)$$

The correction term then becomes

$$\frac{1}{2}\frac{qh^2}{c^2}\cdot\frac{1}{R_s^3}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (40)$$

The orbital velocity then takes the form (for the sun)

$$\frac{d\theta_s}{dt}=\frac{h}{r^2}\left(1-\frac{1}{2}\frac{qh^2}{c^2}\cdot\frac{1}{R_s^3}\right)\equiv\frac{1}{T_s}. \quad . \quad . \quad . \quad (41)$$

For the like orbit, situated at the earth's surface at distance R_e, we have

$$\frac{d\theta_e}{dt}=\frac{h}{r^2}\left(1-\frac{1}{2}\frac{qh^2}{c^2}\cdot\frac{1}{R_e^3}\right)\equiv\frac{1}{T_e}. \quad . \quad . \quad . \quad (42)$$

The ratio of the two time periods is consequently

$$\frac{T_e}{T_s}=\frac{1-\frac{1}{2}\frac{qh^2}{c^2}\cdot\frac{1}{R_s^3}}{1-\frac{1}{2}\frac{qh^2}{c^2}\cdot\frac{1}{R_e^3}}\equiv 1+\frac{1}{2}\frac{qh^2}{c^2}\left(\frac{1}{R_e^3}-\frac{1}{R_s^3}\right). \quad . \quad (43)$$

The corresponding Einsteinian formula is, however,

$$\frac{T_e}{T_s}\equiv 1+\frac{1}{2}k\left(\frac{1}{R_e}-\frac{1}{R_s}\right). \quad . \quad . \quad . \quad . \quad (44)$$

It is very evident therefore that, though under the circumstances the Einsteinian Spectral Shift can be taken to be theoretically justified, nevertheless the effect must be much smaller than that predicted by Einstein's methods.

June 1930.

IX. *On the Steady Rotation of an Anchor Ring in a Viscous Liquid.* *By* F. E. Relton, *M.A., B.Sc., Imperial College of Science and Technology* *.

Part I.

Summary

The anchor ring rotates slowly about its axis, and the path of each particle of the liquid is taken as a circle parallel to the equatorial plane. The solution depends upon Laplace's equation for a uniplanar vector, and the analysis is developed partly in cylindrical coordinates and partly in ring coordinates. An expression for a multiple of the velocity is obtained as a cosine series in which the coefficients are Beta and Riemann P-functions. Two solutions are employed, valid in different regions; the one applies to an annulus with a viscous liquid core, the other to a liquid at rest at infinity. The validity of the result is limited by the ignoration of the inertia terms.

1. When using rectangular coordinates the axes of x, y, z are taken as a right-handed system. In using cylindrical coordinates r, z, ϕ the azimuth is measured from the x-plane, and the direction of ϕ increasing is counter-clockwise when the origin is viewed from the positive end of the z-axis. Two "ring" coordinates λ, θ are also employed, and they are defined as follows: two points A, B being taken on the y-axis at distances $\pm b$ from the origin, for any point P in the x-plane we have $\exp \lambda = \mathrm{PA}/\mathrm{PB}$ and the angle $\mathrm{APB} = \pi - \theta$. Constant values of λ accordingly give a family of coaxal circles with real limiting points at A, B. The values $\exp \lambda = 0, \infty, 1$ correspond to A, B and the z-axis respectively. Constant values of θ give the conjugate family of circles with real intersections at A, B; and as the two families are orthogonal we have in λ, θ, ϕ a system of orthogonal curvilinear coordinates. When the point P traverses the z-axis through the z-values $\infty, 0, -\infty$ the angle θ ranges through $\pi, 0, -\pi$. The anchor ring is generated by rotating about the z-axis the circle defined by $\lambda = c$, the points A, B accompanying.

* Communicated by Prof. S. Chapman, F.R.S.

2. Consideration of the magnitude and direction of PA, PB gives

$$\exp(\lambda - i\theta) = (b - r - iz)/(b + r + iz),$$

whence

$$r + iz = b\{1 - \exp(\lambda - i\theta)\}/\{1 + \exp(\lambda - i\theta)\}.$$

By separation of real and imaginary parts we have

$$r(1 + 2e^{\lambda}\cos\theta + e^{2\lambda}) = b(1 - e^{2\lambda}). \quad . \quad . \quad (2.1)$$

Also

$$dr + i\,dz = 2b(i\,d\theta - d\lambda)\exp(\lambda - i\theta)/\{1 + \exp(\lambda - i\theta)\}^2,$$

whence

$$dr^2 + dz^2 = r^2(d\lambda^2 + d\theta^2)\operatorname{cosech}^2\lambda.$$

For the arc-element we thus have

$$ds^2 = dx^2 + dy^2 + dz^2 = dr^2 + dz^2 + (r\,d\phi)^2$$
$$= r^2\{(d\lambda^2 + d\theta^2)\operatorname{cosech}^2\lambda + d\phi^2\}.$$

For brevity we shall denote $\operatorname{cosech}\lambda$ by h.

3. At any point we take a mutually orthogonal and right-handed system of unit vectors $\mathbf{l}, \mathbf{m}, \mathbf{n}$ in the directions of increasing λ, θ, ϕ respectively. We then have the differential operator

$$\nabla = \frac{1}{rh}\left\{\mathbf{l}\frac{\partial}{\partial\lambda} + \mathbf{m}\frac{\partial}{\partial\theta} + h\mathbf{n}\frac{\partial}{\partial\phi}\right\},$$

and for Laplace's operator

$$\nabla^2 = \frac{1}{r^3h^2}\left\{\frac{\partial}{\partial\lambda}\left(r\frac{\partial}{\partial\lambda}\right) + \frac{\partial}{\partial\theta}\left(r\frac{\partial}{\partial\theta}\right) + \frac{\partial}{\partial\phi}\left(rh^2\frac{\partial}{\partial\phi}\right)\right\}.$$

4. In the absence of body forces, the steady motion of an incompressible liquid, of density ρ and kinematical viscosity ν, is given by the equation

$$\nu\nabla^2\mathbf{v} - \nabla p/\rho = \mathbf{v}\,.\,\nabla\mathbf{v}, \quad . \quad . \quad . \quad (4.1)$$

where $\mathbf{v}$ is the velocity and p is the pressure. In the present problem the "steady motion" will be construed as one in which each particle describes a circle about the z-axis at a uniform rate, so that $\mathbf{v} = \mathbf{n}v$ with $\partial v/\partial\phi = 0$. This is the assumption made by Kirchhoff in the analogous case of the rotating sphere; its adequacy will be mentioned later.

5. With these assumptions we have

$$\mathbf{v}\,.\,\nabla = vr^{-1}\partial/\partial\phi$$

and

$$\mathbf{v}\,.\,\nabla\mathbf{v} = vr^{-1}\partial(\mathbf{n}v)/\partial\phi = v^2r^{-1}\partial\mathbf{n}/\partial\phi.$$

But $\partial\mathbf{n}/\partial\phi$ is evidently perpendicular to the z-axis, and is therefore expressible in terms of $\mathbf{l}$ and $\mathbf{m}$. We shall not trouble to evaluate it here, as its explicit form will not be required for a reason that will appear later ; but we note parenthetically that

$$\partial^2\mathbf{n}/\partial\phi^2 = -\mathbf{n}.$$

Proceeding to the evaluation of $\nabla^2\mathbf{v}$, we have

$$\frac{\partial\mathbf{n}}{\partial\lambda} = 0 = \frac{\partial\mathbf{n}}{\partial\theta},$$

$$\frac{\partial}{\partial\lambda}\left(r\frac{\partial\mathbf{v}}{\partial\lambda}\right) = \frac{\partial}{\partial\lambda}\left(r\frac{\partial\mathbf{n}v}{\partial\lambda}\right) = \mathbf{n}\frac{\partial}{\partial\lambda}\left(r\frac{\partial v}{\partial\lambda}\right),$$

$$\frac{\partial}{\partial\theta}\left(r\frac{\partial\mathbf{v}}{\partial\theta}\right) = \frac{\partial}{\partial\theta}\left(r\frac{\partial\mathbf{n}v}{\partial\theta}\right) = \mathbf{n}\frac{\partial}{\partial\theta}\left(r\frac{\partial v}{\partial\theta}\right),$$

$$\frac{\partial}{\partial\phi}\left(rh^2\frac{\partial\mathbf{v}}{\partial\phi}\right) = rh^2\frac{\partial^2\mathbf{n}v}{\partial\phi^2} = -\mathbf{n}rh^2v\ ;$$

$$\therefore\quad \nabla^2\mathbf{v} = \frac{\mathbf{n}}{r^3h^2}\left\{\frac{\partial}{\partial\lambda}\left(r\frac{\partial v}{\partial\lambda}\right) + \frac{\partial}{\partial\theta}\left(r\frac{\partial v}{\partial\theta}\right) - rh^2v\right\}.$$

6. Adverting to the equation of motion (4.1), symmetry dictates $\partial p/\partial\phi = 0$, and as $\mathbf{v}.\nabla\mathbf{v}$ is expressible in terms of $\mathbf{l}$ and $\mathbf{m}$, the problem depends for its solution on the equation $\nabla^2\mathbf{v} = 0$. A more tractable form can be given to this if we remember that v is a function of position that happens to be independent of ϕ ; for the equation is then equivalent to $\nabla^2 v = v/r^2$ in ring coordinates, with $\partial v/\partial\phi = 0$. Now the operator ∇^2 is invariant, so that the equation can be equally well expressed as $r^2\sqrt{r}\nabla^2 v = v\sqrt{r}$ in cylindrical coordinates, with $\partial v/\partial\phi = 0$. If we utilize the well-known result

$$\sqrt{r}\nabla^2 v = \frac{\partial^2 w}{\partial r^2} + \frac{\partial^2 w}{\partial z^2} + \frac{1}{r^2}\frac{\partial^2 w}{\partial\phi^2} + \frac{w}{4r^2},$$

wherein $w = v\sqrt{r}$, we see that our problem comes to depend upon

$$\frac{\partial^2 w}{\partial r^2} + \frac{\partial^2 w}{\partial z^2} = \frac{3w}{4r^2}. \qquad . \quad . \quad . \quad (6.1)$$

7. Although this last form of the equation is very simple and is soluble in terms of $\exp(\kappa z)r^{\frac{1}{2}}\mathrm{J}_1(\kappa r)$, it is not the most suitable ; for the boundary of the ring is expressed by $\lambda = c$, and not by a sufficiently simple relation between r and z. As w is a function of only two coordinates, we can regard

the coordinates r, z, λ, θ as confined to the x-plane, with (6.1) as a degenerate form of

$$\nabla^2 w = 3w/4r^2, \quad \text{wherein} \quad \partial w/\partial x = 0,$$

the coordinates being r, z, x. As ∇^2 is invariant, the same form of the equation obtains when the coordinates are λ, θ, x. But with these coordinates we have for the arc-element

$$ds^2 = (rh\,d\lambda)^2 + (rh\,d\theta)^2 + dx^2,$$

so that

$$\nabla^2 = \frac{1}{r^2h^2}\left\{\frac{\partial^2}{\partial\lambda^2} + \frac{\partial^2}{\partial\theta^2} + \frac{\partial}{\partial x}\left(r^2h^2\frac{\partial}{\partial x}\right)\right\}.$$

From this it follows that (6·1) is equivalent to

$$\frac{\partial^2 w}{\partial\lambda^2} + \frac{\partial^2 w}{\partial\theta^2} = \frac{3}{4}h^2 w = \frac{3}{4}w\,\mathrm{cosech}^2\lambda.$$

8. For the solution of this last equation we adopt the substitution

$$w = \mathrm{LH}, \quad \text{where} \quad \mathrm{L} = f(\lambda) \quad \text{and} \quad \mathrm{H} = (\mathrm{F})\theta,$$

and so derive

$$\mathrm{L}^{-1}\partial^2\mathrm{L}/\partial\lambda^2 + \mathrm{H}^{-1}\partial^2\mathrm{H}/\partial\theta^2 = (3/4)\,\mathrm{cosech}^2\lambda,$$

which necessitates

$$\partial^2\mathrm{H}/\partial\theta^2 + k^2\mathrm{H} = 0$$

and

$$\partial^2\mathrm{L}/\partial\lambda^2 = \mathrm{L}(k^2 + \tfrac{3}{4}\,\mathrm{cosech}^2\lambda),$$

where k is some real or imaginary constant. Taking k to be real, the H-equation gives

$$\mathrm{H} = \mathrm{A}\cos\kappa\theta + \mathrm{B}\sin\kappa\theta,$$

where A, B are arbitrary constants.

9. The L-equation we transform by the substitution $\exp 2\lambda = t$, whereupon it assumes the Fuchsian type

$$\frac{d^2\mathrm{L}}{dt^2} + \frac{1}{t}\frac{d\mathrm{L}}{dt} - \frac{\mathrm{L}}{4t^2}\left\{k^2 + \frac{3t}{(1-t)^2}\right\} = 0. \quad . \quad (9.1)$$

This obviously has singularities at 0, 1; it is easily verified that there is a third at ∞ and that all three are regular. The corresponding indicial equations give as the respective exponents $\pm\frac{1}{2}k$; $\frac{1}{2}\pm 1$; $\pm\frac{1}{2}\kappa$; so that the solution in the form of Riemann's P-function is

$$\mathrm{L} = \mathrm{P}\left\{\begin{matrix} 0 & \infty & 1 & \\ \frac{1}{2}\kappa & \frac{1}{2}\kappa & 1\frac{1}{2} & t \\ -\frac{1}{2}\kappa & -\frac{1}{2}\kappa & -\frac{1}{2} & \end{matrix}\right\}.$$

Since two pairs of exponents are identical, it follows from a well-known theorem, due to Riemann, that the result is easily expressible in a great variety of forms *.

10. The above P-function can be written as

$$t^{\frac{1}{2}\kappa}(1-t)^{\frac{3}{2}}\,\mathrm{P}\left\{\begin{matrix} 0 & \infty & 1 & \\ 0 & \kappa+1\frac{1}{2} & 0 & t \\ -\kappa & 1\frac{1}{2} & -2 & \end{matrix}\right\},$$

so that a solution of (9.1) is

$$\mathrm{L} = t^{\frac{1}{2}\kappa}(1-t)^{\frac{3}{2}}\mathrm{F}\,(\kappa+1\tfrac{1}{2},\ 1\tfrac{1}{2};\ \kappa+1;\ t).$$

We shall denote this hypergeometric function by $\mathrm{F}_1(\kappa, \lambda)$. If we assume that the annulus rotates about the z-axis the sine terms will not be required in the expression for H, and κ can take all positive integral values, including zero. The number corresponding to γ in $\mathrm{F}(\alpha, \beta; \gamma; t)$ is then a positive integer, so that the second, independent solution of (9.1) would be the somewhat intractable form involving the logarithmic term.

11. The solution involving $\mathrm{F}_1(\kappa, \lambda)$ relates to the singularity $t=0$, and is convergent so long as $t = \exp 2\lambda < 1$; it is therefore suitable when the annulus has a viscous liquid core, and this case we shall consider first. We have

$$w = v\sqrt{r} = \Sigma_\kappa \mathrm{A}_\kappa t^{\frac{1}{2}\kappa}(1-t)^{\frac{3}{2}}\,\mathrm{F}_1(\kappa, \lambda)\cos\kappa\theta, \qquad (11.1)$$

wherein the A_κ are constants, to be determined from the boundary condition of no slip at the surface of the annulus, *i. e.*, $w = \omega r^{\frac{3}{2}}$ when $\lambda = c$ and ω is the angular velocity of the ring

12. In partial analogy with $t = \exp 2\lambda$ we shall write $\tau = \exp c$, and if we let r_c denote the value of r when $\lambda = c$, we have from (2.1)

$$r_c(1+2\tau\cos\theta+\tau^2) = b(1-\tau^2);$$

the boundary condition for no slip becomes

$$\omega b^{\frac{3}{2}}(1+2\tau\cos\theta+\tau^2)^{-\frac{3}{2}} = \Sigma_\kappa \mathrm{A}_\kappa \tau^\kappa \mathrm{F}_1(\kappa, c)\cos\kappa\theta. \quad (12.1)$$

Let us write

$$(1+2\tau\cos\theta+\tau^2)^{-\frac{3}{2}} = \tfrac{1}{2}c_0 + c_1\cos\theta + c_2\cos 2\theta + \ldots + c_\kappa\cos\kappa\theta + \ldots,$$

* Whittaker and Watson, 'Modern Analysis,' 2nd ed. p. 203.

so that the A_κ are determined from

$$A_\kappa \tau^\kappa F_1(\kappa, c)/\omega b^{\frac{3}{2}} = C_\kappa = \frac{1}{\pi}\int_{-\pi}^{\pi} \frac{\cos\kappa\theta \,.\, d\theta}{(1+2\tau\cos\theta+\tau^2)^{\frac{3}{2}}}.$$

The value of this definite integral can be shown to be

$$4(-\tau)^\kappa F(\kappa+1\tfrac{1}{2}, 1\tfrac{1}{2}\,;\ \kappa+1\,;\ \tau^2)/B(\kappa+1, \tfrac{1}{2}),$$

in which the hypergeometric function is what we have agreed to denote by $F_1(\kappa, c)$. A short, direct proof of this result seems difficult to acquire; it is readily deduced from known results due to Gauss*. It appears from this that the values of the coefficients A_κ are independent of the particular value $\lambda = c$. This was more or less to be expected. The physical interpretation is that in the steady state all particles have the angular velocity of the annulus, and rotate with it as if the core were solid.

13. The equation (9.1) has the alternative solution

$$P\left\{\begin{matrix} 0 & \infty & 1 & \\ 1\frac{1}{2} & \frac{1}{2}\kappa & \frac{1}{2}\kappa & (1-t) \\ -\frac{1}{2} & -\frac{1}{2}\kappa & -\frac{1}{2}\kappa & \end{matrix}\right\},$$

which can be written as

$$t^{\frac{1}{2}\kappa}(1-t)^{\frac{3}{2}} P\left\{\begin{matrix} 0 & \infty & 1 & \\ 0 & \kappa+1\frac{1}{2} & 0 & (1-t) \\ -2 & 1\frac{1}{2} & -\kappa & \end{matrix}\right\}$$

$$= t^{\frac{1}{2}\kappa}(1-t)^{\frac{3}{2}} F(\kappa+1\tfrac{1}{2}, 1\tfrac{1}{2}\,;\ 3\,;\ 1-t).$$

Let us denote this hypergeometric function by $F_2(\kappa, \lambda)$. The solution relates to the singularity $t=1$, and is convergent so long as $1-t = 1-\exp 2\lambda < 1$; we shall see that it proves suitable for the case of an annulus rotating in an infinite liquid at rest at infinity. In equations (11.1) and (12.1) we have merely to replace F_1 by F_2. The coefficients A_κ are then given as

$$A_\kappa = (-)^\kappa 4\omega b^{\frac{3}{2}} F_1(\kappa, c)/F_2(\kappa, c)\,.\,B(\kappa+1, \tfrac{1}{2}).$$

14. The solution corresponding to (11.1) is now

$$w = v\sqrt{r} = 4\omega b^{\frac{3}{2}}(1-t)^{\frac{3}{2}}\Sigma_\kappa(-t^{\frac{1}{2}})^\kappa$$
$$\{F_1(\kappa, c)\,F_2(\kappa, \lambda)/F_2(\kappa, c)\,B(\kappa+1, \tfrac{1}{2})\}\cos\kappa\theta. \quad (14.1)$$

* Forsyth's Diff. Eqns., 6th ed. p. 240.

This satisfies the condition of no slip at the surface of the annulus; but it remains to examine its behaviour on the z-axis and at infinity, where the liquid should be at rest. On the z-axis and at infinity we have

$$\exp 2\lambda = 1 = t, \quad \text{so that} \qquad F_2(\kappa, \lambda) = 1.$$

At infinity we have additionally $\theta = \pi$, and although a zero factor now precedes the Σ in (14.1), we ought to examine the series

$$\Sigma_\kappa F_1(\kappa, c)/F_2(\kappa, c)\, B(\kappa+1, \tfrac{1}{2}), \quad . \quad . \quad (14.2)$$

and this we shall now do by considering separately the factors in the general term.

15. Utilizing the asymptotic formula

$$\Gamma(x) \to x^x e^{-x}(2\pi/x)^{\frac{1}{2}}(1 + \ldots),$$

we find that, with increasing κ,

$$B(\kappa+1, \tfrac{1}{2}) \to \pi^{\frac{1}{2}}/(\kappa+1)^{\frac{1}{2}}.$$

The expansion of $F_1(\kappa, c)$ is term by term less than the expansion of

$$\{1-\tau^2(\kappa+1\tfrac{1}{2})/(\kappa+1)\}^{-\frac{3}{2}},$$

so that if α be a constant such that $\tau^{-2} > \alpha > 1$, we must have

$$F_1(\kappa, c) < (1-\alpha\tau^2)^{-\frac{3}{2}}$$

for sufficiently large values of κ. The expansion of $F_2(\kappa, c)$ is term by term greater than the expansion of

$$\{1-\tfrac{1}{2}(1-\tau^2)\}^{-(\kappa+1\frac{1}{2})} = \left(\frac{2}{1+\tau^2}\right)^{\kappa+1\frac{1}{2}},$$

whence we conclude that, to a constant factor près, the general term of (14.2) is ultimately less than

$$(\kappa+1)^{\frac{1}{2}}\{\tfrac{1}{2}(1+\tau^2)\}^{\kappa+1\frac{1}{2}} = (\beta n)^{\frac{1}{2}}\beta^n,$$

where

$$n = 1+\kappa \quad \text{and} \quad \beta = \tfrac{1}{2}(1+\tau^2) < 1.$$

The series whose general term is $n\beta^n$ is convergent; we conclude that (14.2) is convergent, whence (14.1) converges to zero when $\exp\lambda = 1$, and equally so when $\theta = \pi$.

16. Adverting to the equation of motion (4.1), the pressure is to be determined from $\mathbf{v}.\nabla\mathbf{v} = -\nabla p/\rho$, and our choice of $\mathbf{v} = n\mathbf{v}$ makes $\mathbf{v}.\nabla\mathbf{v}$ a vector perpendicular to the z-axis. Accordingly the level pressure surfaces would be circular cylinders about the z-axis; but the attempt to evaluate p is

frustrated by the fact that the expression for dp, derived from (4.1), turns out to be not an exact differential. We accordingly fall back on the trivial solution $p = \text{constant}$, with the customary ignoration of the inertia terms. The explanation of the anomaly, such as it is, has already been given by Stokes in the case of the rotating sphere. The body acts as a centrifuge and the assumption $\mathbf{v} = \mathbf{n}v$ for a steady motion is not strictly tenable. A circulatory motion in the meridian plane should be superimposed.

Part II.

Summary.

In the present part of the paper the self-conjugate "rate of strain" dyadic is shown to consist of a single shear whose surface-value can be computed from the value of the velocity previously obtained. The surface value of this shear leads to an expression for the resisting couple.

---

17. We begin by finding some differential coefficients of the unit vectors with respect to the coordinates. Since $\mathbf{n}$ is always perpendicular to the meridian plane, we have

$$\partial\mathbf{n}/\partial\theta = 0 = \partial\mathbf{n}/\partial\lambda. \quad . \quad . \quad . \quad . \quad (17.1)$$

By considering the elementary rectangle, two of whose adjacent sides are the arc-elements $\mathbf{l}rh\,d\lambda$, $\mathbf{n}r d\phi$, we have

$$\frac{\partial}{\partial\phi}(\mathbf{l}rh\,d\lambda)\,d\phi = \frac{\partial}{\partial\lambda}(\mathbf{n}r d\phi)\,d\lambda,$$

which in virtue of (17.1) gives

$$rh\,\partial\mathbf{l}/\partial\phi = \mathbf{n}\,\partial r/\partial\lambda. \quad . \quad . \quad . \quad . \quad (17.2)$$

From similar considerations, or more simply by replacing $\mathbf{l}$ and λ by $\mathbf{m}$ and θ respectively, we derive

$$rh\,\partial\mathbf{m}/\partial\phi = \mathbf{n}\,\partial r/\partial\theta. \quad . \quad . \quad . \quad . \quad (17.3)$$

18. The condition of perpendicularity $\mathbf{l}.\mathbf{n} = 0$ yields

$$\mathbf{l}.\partial\mathbf{n}/\partial\phi + \mathbf{n}.\partial\mathbf{l}/\partial\phi = 0, \quad . \quad . \quad . \quad (18.1)$$

and a similar equation holds, wherein $\mathbf{m}$ replaces $\mathbf{l}$. The vector $\partial\mathbf{n}/\partial\phi$, being perpendicular to the z-axis, is expressible in terms of $\mathbf{l}$ and $\mathbf{m}$; we have, in fact,

$$\begin{aligned}\partial\mathbf{n}/\partial\phi &\equiv \mathbf{l}\,\mathbf{l}.\partial\mathbf{n}/\partial\phi + \mathbf{m}\,\mathbf{m}.\partial\mathbf{n}/\partial\phi \\ &= -\mathbf{l}\,\mathbf{n}.\partial\mathbf{l}/\partial\phi - \mathbf{m}\,\mathbf{n}.\partial\mathbf{m}/\partial\phi \text{ by (18.1)} \\ &= -(\mathbf{l}\,\partial r/\partial\lambda + \mathbf{m}\,\partial r/\partial\theta)/rh, \quad . \quad . \quad . \quad (18.2)\end{aligned}$$

this last in virtue of (17.2) and (17.3).

19. The self-conjugate "rate of strain" dyadic is

$$\Phi = \tfrac{1}{2}\{\nabla \mathbf{v} + \mathbf{v}\nabla\}, \quad . \quad . \quad . \quad . \quad (19.1)$$

where the velocity $\mathbf{v} = \mathbf{n}v$, with $\partial v/\partial\phi = 0$; the differential operator

$$\nabla = (rh)^{-1}\{\mathbf{l}\partial/\partial\lambda + \mathbf{m}\,\partial/\partial\theta + h\mathbf{n}\,\partial/\partial\phi\}.$$

When one of the dyadics in (19.1) has been calculated, the other can be written down by simply reversing the order of the vectors in each of the resulting dyads. Thus we have

$$\begin{aligned}\nabla\mathbf{v} &= (rh)^{-1}\{\mathbf{l}\partial/\partial\lambda + \mathbf{m}\,\partial/\partial\theta + h\mathbf{n}\,\partial/\partial\phi\}\mathbf{n}v \\ &= (rh)^{-1}\{\mathbf{ln}\,\partial v/\partial\lambda + \mathbf{mn}\,\partial v/\partial\theta + vh\mathbf{n}\,\partial\mathbf{n}/\partial\phi\} \text{ by (17·1)} \\ &= (rh)^{-1}\{\mathbf{ln}\,\partial v/\partial\lambda - \mathbf{nl}\,v\partial r/r\partial\lambda + \mathbf{mn}\,\partial v/\partial\theta \\ &\qquad - \mathbf{nm}\,v\partial r/r\partial\theta\} \text{ by (18.2)},\end{aligned}$$

whence it follows that (19.1) becomes

$$\begin{aligned}2\Phi &= (\mathbf{ln}+\mathbf{nl})\left(\frac{1}{rh}\frac{\partial v}{\partial\lambda} - \frac{v}{r^2h}\frac{\partial r}{\partial\lambda}\right) \\ &\qquad + (\mathbf{mn}+\mathbf{nm})\left(\frac{1}{rh}\frac{\partial v}{\partial\theta} - \frac{v}{r^2h}\frac{\partial r}{\partial\theta}\right) \\ &= (\mathbf{ln}+\mathbf{nl})\frac{1}{h}\frac{\partial}{\partial\lambda}\left(\frac{v}{r}\right) + (\mathbf{mn}+\mathbf{nm})\frac{1}{h}\frac{\partial}{\partial\theta}\left(\frac{v}{r}\right) \\ &= (\mathbf{ln}+\mathbf{nl})h^{-1}\,\partial\dot{\phi}/\partial\lambda + (\mathbf{mn}+\mathbf{nm})h^{-1}\,\partial\dot{\phi}/\partial\theta, \quad (19.2)\end{aligned}$$

where $\dot{\phi}$ is the angular velocity about z-axis.

20. At the surface of the annulus the angular velocity of the liquid is constant, and since the annulus is defined by $\lambda = c$ it appears that the surface value of the **mn** dyadic in (19.2) is zero. Thus, as might have been expected, the surface value of the "rate of strain" dyadic reduces to a single shear in the **ln** plane. The resisting couple, using surface values, is

$$\begin{aligned}&\mu\int_{-\pi}^{\pi} h^{-1}(\partial\dot{\phi}/\partial\lambda)\,2\pi r\,.\,r\,.\,rh\,d\theta \\ &= 4\pi\mu\int_0^{\pi} r^3(\partial\dot{\phi}/\partial\lambda)\,\partial\theta. \quad . \quad . \quad . \quad . \quad . \quad . \quad (20.1)\end{aligned}$$

The evaluation of this integral necessitates using the velocity, and as an expression has been obtained for

$$w = vr^{\frac{1}{2}} = \dot{\phi}r^{\frac{3}{2}},$$

we can write the integrand

$$r^3(\partial\dot{\phi}/\partial\lambda) = r^{\frac{3}{2}}\,\partial w/\partial\lambda - 1\tfrac{1}{2}r^2\omega\,\partial r/\partial\lambda. \quad . \quad (20.2)$$

21. It is an advantage to choose a new independent variable $\rho = \exp\lambda$, so that

$$\partial/\partial\lambda = \rho\partial/\partial\rho \quad \text{and} \quad r = b(1-\rho^2)/(1+2\rho\cos\theta+\rho^2).$$

From these we deduce, after a little reduction, that

$$r^2\partial r/\partial\lambda = b^3(1-\rho^2)^2\{(1-\rho^2)^2\zeta^4-(1+\rho^2)\zeta^3\}, \quad (21.1)$$

where

$$\zeta = (1+2\rho\cos\theta+\rho^2)^{-1}.$$

By integrating (21.1) we have *

$$\int_0^\pi r^2(\partial r/\partial\lambda)\,.\,d\theta = \pi b^3(1-\rho^2)^2\{(1-\rho^2)^2\,\mathrm{F}(4,4;1;\rho^2) - (1+\rho^2)\,\mathrm{F}(3,3;1;\rho^2)\}.$$

At the surface of the annulus we put $\exp c = \tau$, so that the contribution to the resisting couple made by the last term in (20.2) is

$$6\pi^2\mu\omega b^3(1-\tau^2)^2\{(1+\tau^2)\,\mathrm{F}(3,3;1;\tau^2) - (1-\tau^2)^2\,\mathrm{F}(4,4;1;\tau^2)\}. \quad . \quad . \quad (21{\cdot}2)$$

22. In order to evaluate the contribution made by the first term on the right in (20.2), let us take w_κ as the general term in the expansion of w, so that from (14.1)

$$w_\kappa \doteq (-\rho)^\kappa(1-\rho^2)^{\frac{3}{2}}\,\mathrm{F}_1(\kappa,c)\,\mathrm{F}_2(\kappa,\lambda)/\mathrm{F}_2(\kappa,c)\,\mathrm{B}(\kappa+1,\tfrac{1}{2}),$$

where irrelevant factors have been omitted. We then have

$$\frac{\partial w_\kappa}{\partial\lambda} = \rho\frac{\partial w_\kappa}{\partial\rho} = w_\kappa\left[\frac{\kappa-\rho^2(\kappa+3)}{1-\rho^2} - \rho^2(\kappa+1\tfrac{1}{2})\frac{\mathrm{F}(\kappa+2\tfrac{1}{2},2\tfrac{1}{2};4;1-\rho^2)}{\mathrm{F}(\kappa+1\tfrac{1}{2},1\tfrac{1}{2};3;1-\rho^2)}\right]. \quad (22.1)$$

For brevity we will let H_κ denote the surface value of $[\ ]$ in (22·1), obtained by simply replacing ρ by τ. The surface value of r is

$$b(1-\tau^2)/(1+2\tau\cos\theta+\tau^2),$$

and the surface value of w_κ is

$$4\omega b^3(-\tau)^\kappa(1-\tau^2)^{\frac{3}{2}}\cos\kappa\theta\,.\,\mathrm{F}(\kappa+1\tfrac{1}{2},1\tfrac{1}{2};\kappa+1;\tau^2)/\mathrm{B}(\kappa+1,\tfrac{1}{2}). \quad . \quad . \quad (22.2)$$

* *Cf.* Forsyth, *l. c.*

Utilizing the result

$$\int_0^{\pi} (1+2\tau\cos\theta+\tau^2)^{-\frac{3}{2}}\cos\kappa\theta\,.\,d\theta$$

$$= 2\pi(-\tau)^{\kappa}\mathrm{F}(\kappa+1\tfrac{1}{2},\ 1\tfrac{1}{2};\ \kappa+1;\ \tau^2)/\mathrm{B}(\kappa+1,\ \tfrac{1}{2}),$$

we have that, over the surface,

$$\int_0^{\pi} r^{\frac{3}{2}}\,(\partial w_{\kappa}/\partial\lambda)\,\partial\theta = 8\pi\omega b^3\mathrm{H}_{\kappa}\tau^{2\kappa}(1-\tau^2)^3$$

$$\{\mathrm{F}(\kappa+1\tfrac{1}{2},\ 1\tfrac{1}{2};\ \kappa+1\ ;\ \tau^2)/\mathrm{B}(\kappa+1,\ \tfrac{1}{2})\}^2. \quad . \quad (22.3)$$

We conclude that the resisting couple is the expression (21.2), together with

$$32\pi^2\mu\omega b^3(1-\tau^2)^3\sum_{\kappa=0}^{\infty}\mathrm{H}_{\kappa}$$

$$\{\tau^{\kappa}\,\mathrm{F}(\kappa+1\tfrac{1}{2},\ 1\tfrac{1}{2};\ \kappa+1;\ \tau^2)/\mathrm{B}(\kappa+1,\ \tfrac{1}{2})\}^2.$$

23. In conclusion we may comment on the persistent use of the hypergeometric function throughout both parts of this paper. It was remarked in § 9 that the Riemann P-function was capable of many forms in virtue of the identity of two pairs of exponents. If comparison be made with the P-function expressing the solution of Legendre's associated equation, it becomes evident that our analysis might have been conducted in terms of associated Legendre functions of degree half an odd integer. In some ways this would have been a gain, especially if the argument were an even hyperbolic function of λ; for the ring which has been defined by $\lambda = c$ is equally well defined by $\lambda = -c$. Such a solution would have been more in line with the known solution of Laplace's equation in ring coordinates, and, if need be, the change from hypergeometric functions to associated Legendre functions could be made posteriorly by means of Barnes's formulæ*. The change offers no advantages for computative purposes; there is a dearth of tabulated values of associated Legendre functions of fractional degree, and though asymptotic formulæ are available† where large numbers are concerned, they discriminate but little between the hypergeometric function and the associated Legendre function. And they afford no help in the earlier terms of the series. The adherence to hypergeometric functions was dictated by the availability of Gauss's results, giving powers of our r in Fourier series, wherein the coefficients are hypergeometric functions.

* Bateman, 'Wave Motion,' p. 106.

† Watson, Trans. Camb. Phil. Soc. xxii. p. 277 (1918).

X. *Photophoresis and the Influence upon it of Electric and Magnetic Fields. By* Felix Ehrenhaft, *Dr. Phil., Acting Professor of Physics and Head of the Physics Institute III. of the University of Vienna* *.

IN the year 1910† I had already endeavoured to work out my method of determination of electrical charges on single submicroscopic particles, suspended in my electrical condenser, by removing all the possible accessory circumstances which could disturb the measurements. I then also adjusted the illuminating light so as to be able to observe the light pressure on these small particles. In this connexion I observed that single submicroscopic particles which were falling with uniform velocity in a gas as frictional medium were rapidly carried away in the horizontal direction, when they came into the intense part of the illuminating beam of light. After passing the intense beam, these particles again fell perpendicularly. Even at that time I ascribed this striking phenomenon to light pressure, or to the radiometric forces which become observable on small particles of a definite size. This phenomenon was first observed on metals like silver and gold, and then also on particles of selenium and tellurium.

Later I submitted this phenomenon to a special examination, and found that the particles of various substances move in the same beam of light in opposite directions simultaneously. I have named this phenomenon " Photophoresis " ‡, and designated the movement in the direction of the light propagation as positive photophoresis, the opposite movement as negative photophoresis. I then examined whether photophoresis depends upon the pressure of the surrounding gas. It was found that this phenomenon is to a first approximation independent of the gas pressure, at least in the interval from atmospheric pressure to about 300 mm. Hg. This fact was striking with regard to the intervention of radiometric forces in the explanation of this phenomenon, since, as is well known, they depend upon the pressure.

* Communicated by the Author.

† F. Ehrenhaft, *Sitzungsber. d. D. Akad. d. Wiss.*, Vienna, May 12th, 1910, p. 835.

‡ F. Ehrenhaft, *Ann. d. Phys.* lvi. p. 81 (1918).

In the course of further investigations I found that, besides the movement of the particles in the direction of the light propagation, or in the opposite direction (longitudinal effect), the light also imparts impulses upon matter perpendicularly to those directions (transverse effect), if the particle is acted on by an electric field transversely to the light propagation *. In spite of the fact, proved experimentally, that this phenomenon is independent of the gas pressure in the interval mentioned above—a fact which could be understood in the case of the light pressure on small spheres—the assumption was plausible that radiometric forces are responsible for these phenomena, for a negative light pressure would be in contradiction to the impulse law of electrodynamics.

The problem of the distribution of the electromagnetic energy around an absolutely conducting sphere (J. J. Thomson, F. Ehrenhaft †) had already been treated. G. Mie ‡ has calculated this problem for spheres of any material (complex refractive index). Schwarzschild has calculated the light pressure on black or on perfectly reflecting spheres, and P. Debye § the light pressure on spheres of any material.

The attempt to set up a theory of the phenomena, described above, was made by A. Rubinowicz ‖, G. Laski and F. Zerner ¶, and G. Hettner ** in the following manner. The energy of the light wave falling on the small particle will be partly absorbed by it. It is possible to calculate the field of the light wave within the particle; further, the Poynting vector can be calculated, and consequently the space distribution of the heat-sources within the sphere. The distribution of the temperature within the sphere and in the surrounding gas can then be obtained by fixing the boundary conditions, which are essentially determined by the conduction of heat between the particle and the gas, and by the continuity of the

* F. Ehrenhaft and D. Konstantinowsky, *Anzeiger der Wiener Akademie der Wissenschaften*, No. 9 (18 March, 1920).

† J. J. Thomson, Rec. Res. p. 361; F. Ehrenhatt, *Sitzungsber. d. Wiener Akad. d. Wiss.* cxii. p. 232 (1903).

‡ G. Mie, *Ann. d. Phys.* [4] xxv. p. 377 (1908).

§ P. Debye, *Ann. d. Phys.* [4] xxx. p. 37 (1909).

‖ A. Rubinowicz, *Ann. d. Phys.* lxii. pp. 691 and 716 (1920); *ZS. für Phys.* vi. p. 405 (1921).

¶ G. Laski and Zerner, *ZS. für Phys.* iii. p. 224 (1920); vi. p. 411 (1921).

** G. Hettner, *ZS. für Phys.* xlvii. p. 499 (1928).

heat-flow. In this connexion the two following cases will be discussed separately.

1. The illuminated sphere may be small in comparison with the mean free path of the molecules. The collisions in the neighbourhood of the spherical surface will then not be changed by the sphere ; the gas will accordingly have the same velocity distribution, and therefore no gradient of temperature. This fact can also be expressed as follows :—The heat-conduction of the gas is infinitely large ; there is accordingly a jump of temperature at the surface of the sphere with respect to the gas.

2. In the other case the particle is large in comparison with the mean free path of the molecules. Here the velocity distribution of the molecules will be changed by the sphere ; the distribution of temperature in the gas and in the sphere will not only be determined by the surface of the sphere, but also by the heat-conduction within the gas. There is in this case a distribution of temperature also in the gas.

In both cases, for the purpose of the determination of the radiometric force, the question arises as to the value of the mechanical impulse conferred upon the particle by the gas, in consequence of the jump or of the gradient of temperature. The negative value of this impulse-current would be the photophoretic force on the sphere.

If we take into account the whole experimental and theoretical facts, we must regard it as desirable to explain the remarkable phenomena of photophoresis by further experiments. This may be confirmed by the circumstance that the observed particles are of the order of magnitude of the mean free paths of the molecules ($l \sim a \sim 10^{-5}$ cm.), and that for these conditions the radiometric force has not been calculated. For this interval G. Hettner has set up an interpolation formula, the validity of which still remains to be proved. J. Mattauch * has made use of this formula, as his experiments showed that the photophoresis depends upon the pressure, especially at low pressures. There is further need for an exact determination of the magnitude of the radiometric force and for its comparison with the theory. Also the temperature gradient on the particle, produced by the radiometric force, is not yet known exactly. Moreover, it is remarkable

* J. Mattauch, *Ann. d. Phys.* lxxxv. p. 967 (1928).

that, in spite of the many experiments, it was hitherto not possible to transform a light-negative particle into a light-positive one by evacuation, and this must occur if the radiometric theory be valid and the light pressure be large enough to be observed.

The complication of the whole matter induced me to attack this problem again from the experimental side. As the theory was only interested in exploring longitudinal photophoresis, and paid less attention to the transverse effect, it seemed especially important to me to submit the latter to an examination, and to explore generally the dependence of photophoresis upon electric and magnetic forces. In this connexion certain results were obtained which I, my assistants (E. Wasser and M. Reiss), and my pupils (P. Selner and E. Reeger) will characterize by the following four experiments.

Experiment I.—Transverse Magneto-photophoresis.*

If we illuminate a homogeneous magnetic field† of about 7000 gauss perpendicularly to the direction of the lines of force from one side with an intensive suitably-concentrated beam of light, and from the other side with diffuse light, and, further, arrange to be able to observe one and the same particle at right angles to the illuminating beams by means of a microscope, alternately with the concentrated and the diffused light, it is possible simultaneously to adjust a homogeneous electric field parallel to the homogeneous magnetic field. We can then observe that single particles of nickel, iron, antimony, etc., produced in an electric arc—which remain stationary or fall slowly in the diffused light under the influence of the homogeneous magnetic field—move in the concentrated beam in the direction of the magnetic lines of force, and hence perpendicularly to the illuminating beam. Moreover, this movement is reversed on reversing the magnetic field. When several particles of the same material are in the field of vision of the microscope, one can see that, on establishing the magnetic field, some particles move in the direction of the magnetic field intensity and others in the opposite direction. This effect takes place on some

* By "transverse" and "longitudinal" photophoresis we understand, in what follows, movements perpendicular and parallel respectively to the illuminating beam of light.

† F. Ehrenhaft and E. Wasser, Phil. Mag. v. (Feb. 1928).

particles even in the weak remanent field of the "micro-magnet," and does not increase on these particles when the magnetic intensity is raised.

Experiment II.—Longitudinal Magneto-photophoresis.

In this arrangement, as in experiment I., many particles of Ni and Fe show longitudinal photophoresis, both light-positive (in the direction of the light propagation) and light-negative (towards the light source). On establishing the magnetic field perpendicular to the illuminating beam, and simultaneously compensating a possible transverse effect on charged particles by means of a suitable electric field, the longitudinal photophoresis increases with some particles and diminishes with others. There are also particles which show no photophoresis without the magnetic field, but the photophoresis appears immediately on switching on the field. The longitudinal magneto-photophoresis, here described, cannot be reversed by changing the direction of the magnetic field.

Experiment III.—Transverse Electro-photophoresis.

Suppose we adjust the beams of light in an Ehrenhaft condenser as in experiment I., and remove all charged particles by means of a suitable electrical field, so that in the diffused light no movement can be observed except the free fall of the particles. If we then illuminate these particles with the intense beam of light, they move in the homogeneous electric field along the lines of force, and so perpendicularly to the illuminating beam. Some particles move, as in experiment I., in the direction of the electric force, others in the opposite direction. The movement stops immediately on cutting off the concentrated light. The intensity of the movement depends upon the magnitude and kind of the particle of the same material. It is especially dependent on the material. While with tellurium, antimony, and iodine this movement takes place even for very small electric fields (0·02 e.s.u.), it appears with nickel, iron, selenium, and bismuth only at about 20 e.s.u. The last-mentioned movement with strong electric fields G. Placzek has already seen in my Institute. This transverse electro-photophoresis can, just as in experiment I., almost always be reversed by means of the electric field. The velocity of the movement

described increases with the electric intensity, but is not proportional to it; it grows more slowly than the electric field intensity, and seems to attain a saturation value. In a constant electric field the velocity of the movement increases with the intensity of the light. Simultaneous illumination of the particle on the left and the right causes an increase in the velocity. Magneto-photophoresis, as well as electro-photophoresis, cannot be reversed by changing the direction of the illuminating light.

If the experiment described is made in an alternating electric field (48 periods), there is no transverse electro-photophoresis, even in fields of 20 e.s.u. If, moreover, an alternating field is superposed on the constant electric field, the electro-photophoresis decreases. By these means we were able to inhibit the movement of some particles.

Experiment IV.—Longitudinal Electro-photophoresis.

For the purpose of investigating this case, charged particles were first observed in an alternating field. We found some particles the longitudinal photophoresis of which was not influenced, and some the photophoretic velocity of which was reduced in the alternating field. The further experiments, concerning the photophoresis of uncharged particles under the influence of constant and of alternating electric fields, have shown that the photophoretic velocity of such particles also becomes larger or smaller in the fields mentioned*.

If we adopt the conception that photophoresis is to be explained by means of radiometric forces which arise from the distribution of temperature on the particle, then we must assume that under the action of an electric or magnetic field the original distribution of the temperature is changed. In this case the position of the absorption sources can be changed, and by this the lines of the heat-current within the particle or on its surface. This can lead to electrothermal and thermomagnetic effects. The other possibility is to assume that the conducting power between the particle and the gas or the coefficient of accommodation † are changed in the field.

* There are particles of iron and nickel which show in the arrangement described above a movement perpendicularly to the plane: force—beam of light. This effect manifests itself in the disappearance of the particles out of the field of vision of the microscope.

† M. v. Smoluchowsky, Wied. *Ann.* lxiv. p. 101 (1898).

Similar actions of the magnetic field on metallic plates or on thin films, when carrying a thermal current, are known. Moreover, these effects are, when transverse, proportional to the field-intensity, when longitudinal to the square of it. Analogous electrothermal effects are not yet known. Their existence with large bodies would seem probable according to experiments III. and IV. Of course, we must realize that such effects, according to the electromagnetic theory, can take place only on the surface of the particle. The question might be raised, however, as to how far the electrothermal effect can penetrate the particle.

Finally, it may be noticed that all phenomena can be measured exactly, and therefore all experiments can be settled numerically. For the purpose of explanation of the above experiments it may be recognized that the particles, when not spherical, can only set in the direction of the lines of force ; a movement of these particles can not, however, take place in a homogeneous field. The reversible movement here described can therefore only be traced back to the light current. These phenomena, which are new from the experimental standpoint, will be investigated further in my Institute.

My thanks are due to Dr. W. Lawson for the revision and correction of the style of the manuscript.

XI. *The Application of Least Squares.* *By* W. Edwards Deming, *Ph.D., Associate Physicist, Bureau of Chemistry and Soils, U.S. Department of Agriculture, Washington, D.C.* *

Abstact.

The principle of least squares is applied to the general problem of finding the most probable values of observed quantities and of parameters that are related by equations of various sorts. The entire derivation is based on the statement of the principle that the weighted sum of the squares of the residuals is to be minimized with respect to the most probable values of the observed quantities and of any parameters that

* Communicated by the Author. This paper was read at the New York Meeting of the American Physical Society, February 22, 1930.

enter relations between these quantities. These relations will vary in form and number from one type of problem to another. The general solution takes the form of symmetrical equations, and the term "general normal equations" is suggested for them. For a specific type of problem these reduce to a form characteristic of that type.

The application to curve fitting is discussed in detail because with it there is considerable misunderstanding regarding the purpose and meaning of the principle of least squares. For curve fitting, two sets of symmetrical equations are derived from the general solution. In one the only unknowns are the corrections to the approximate values of the parameters, and the name "p normal equations" is suggested for them. The other set adjusts the observations by determining the residuals. The practice of throwing all the adjustment on to one group of observations and none on to the others sometimes violates the foundation of least squares and yields unreasonable results, which bring about unjustified criticism of the principle. The present solution allows the proper weight to be given to each observation. The case where the curve (or surface) must pass rigorously through certain points is included. The equations are written in full for observations to be fitted by a straight line that must pass through a fixed point. A problem in curve fitting is worked out with various assignments of weights to illustrate the effect of throwing the weight on to the wrong set.

An expression is derived for the direction cosines of the line joining the observed and most probable positions of a point plotted to orthogonal axes. When the observations that represent the coordinates of the observed position of a point are of equal weight, this line is normal to the empirical surface, which represents the locus of the most probable positions of all the points. Attention is called to the fact that least squares professes to find the most probable values of the observed quantities, as well as of the parameters, in a problem of curve fitting.

THIS paper will deal with the application of least squares to a general proposition, and it will then employ the general solution for special types of problems. The term "most probable" will have its mathematical significance; whether this is synonymous with "most nearly right" is another matter, as is also the question whether curve fitting by least squares is ever worth while. These are weighty

questions, but they are entirely irrelevant to a discussion that has for its object a clearer understanding of the application of the method.

Some important deductions from the principle of least squares have appeared, presumably for the first time, within the last decade. It seems strange that the classic writers on the subject were not aware of most of these deductions and also of the results to be derived in this article ; but if they were, then apparently they did not consider them worth writing about, perhaps because they did not foresee the confusion and misunderstanding that has been spread by most (if not all) of the writers of textbooks on least squares. The prevalence of erroneous ideas was pointed out in papers by R. Meldrum Stewart * and by Professor Uhler † several years ago. Portions of the discussions in the present article are light, originally from these two papers, that has been reflected by the writer.

Stewart incidentally formulated a general problem for the application of least squares and indicated a method of solution, but did not have occasion to carry it through. One valuable feature of Uhler's paper was his correct solution for the case of pairs of observations supposed to be related by a linear equation ; this much of his paper will be duplicated in the present article, which will contain some of his results as a special case. By solving the general problem to be proposed shortly, we have at once the application to several types of problems that are special cases of the general one.

We shall deal with observations on different kinds of quantities, for example, observations on the pressure, volume, and temperature of a gas, in which case there are three kinds of measurements. If there are n observations on q different kinds of quantities, they can conveniently be designated by $x_1, x_2, \ldots, x_n$; $x_{n+1}, x_{n+2}, \ldots x_{2n}$; $\ldots$; $\ldots$, x_{qn}. w_i will denote the weight of x_i, and r_i will denote its most probable value. The principle of least squares demands that

$$\phi \equiv \sum_i w_i(r_i - x_i)^2$$

be minimized with respect to the r_i and any parameters that may enter into whatever relations exist between the r_i. If there are no such relations, then the r_i are independent and the principle declares that the weighted arithmetic mean of the observations on a quantity is its most probable value.

* R. Meldrum Stewart, Phil. Mag. xl. p. 217 (1920).
† H. S. Uhler, J. O. S. and R. S. I. vii. p. 1043 (1923).

But if there are relations between the r_i, then ϕ is to be a *constrained* minimum. The transition from the simple case, where the r_i are independent, to the most involved cases where there may be observations both on certain quantities and on functions of them, does not require a restatement of the principle of least squares. The *principle* remains the same for all types of problems; the relations between the r_i constitute the distinction between the various types.

If we have the observations P, V, T for the pressure, volume, and temperature of a gas, there is one extra observation because there is one relation between the pressure, volume, and temperature of a gas, namely, its equation of state. The same situation exists with the measurements on the three angles A, B, C of a plane triangle; there is one extra observation; theoretically, measurements of any two should determine all three, because, again, there is one relation between them, namely, $A+B+C=180^0$. The equation of state involves parameters: the rigorous geometrical condition does not. If the parameters are known, then we can find the most probable values of the pressure, volume, and temperature that were observed to be P, V, T. By taking a series of measurements P_h, V_h, T_h ($h=1, 2, \ldots, n$) on the gas, the method of least squares will determine the most probable values of the parameters *and* of the n pressures, n volumes, and n temperatures that were observed to be P_h, V_h, T_h. Right here, incidentally, we have encountered a peculiarity of the least squares method of curve fitting; it professes not only to find values for the parameters in a formula, but also to correct the observations. This is easily overlooked, since curve fitting is usually done for the specific purpose of finding parameters, not of adjusting the observations. In fact, as far as I know, there has been hitherto no solution available by which the observations *could* be adjusted or the parameters found, correctly, except when one set of observations is alone subject to error (*vide infra*).

The word "empirical" will be used in the sense of empirical or rational, since there need be no distinction as far as this discussion is concerned.

As stated, the r_i will in general be related by m equations of condition:

$$F(r_1, r_2, \ldots, r_{qn};\ a, b, \ldots, p)=0,\quad h=1, 2, \ldots, m. \quad (1)$$

$a, b, \ldots, p$ represent the most probable values of the p parameters that may enter. There may also exist l ($l<p$ and $m+l<qn+p$) relations between the parameters:

$$F^h(a, b, \ldots, p) = 0,\quad h = m+1, m+2, \ldots, m+l. \quad (2)$$

ϕ is to be minimized with respect to $r_1, r_2, \ldots, r_{qn}, a, b, \ldots, p$, which are subject to conditions (1) and (2). We shall assume that approximate values $a_0, b_0, \ldots, p_0$ have been obtained by some means; further that the corrections $a-a_0, b-b_0, \ldots, p-p_0$ and the residuals $r_i - x_i$ are so small that the condition equations can be made linear in them without appreciable error.

Then, by writing

$$\mathrm{F}_0^h \equiv \mathrm{F}^h(x_1, x_2, \ldots, x_{qn}\ ;\ a_0, b_0, \ldots, p_0),$$

$$\mathrm{F}_i^h \equiv \frac{d}{dx_i}\mathrm{F}^h(x_1, x_2, \ldots, x_{qn}\ ;\ a_0, b_0, \ldots, p_0),$$

$$\mathrm{F}_a^h \equiv \frac{d}{da_0}\mathrm{F}^h(x_1, x_2, \ldots, x_{qn}\ ;\ a_0, b_0, \ldots, p_0), \text{ etc.,}$$

the following equations replace (1) and (2):

$$\sum_i \mathrm{F}_i^h(r_i - x_i) + \mathrm{F}_a^h(a-a_0) + \mathrm{F}_b^h(b-b_0) + \ldots + \mathrm{F}_p^h(p-p_0) = -\mathrm{F}_0^h, \quad h = 1, 2, \ldots, m\ ; \quad . \quad . \quad . \quad . \quad (3)$$

$$\mathrm{F}_a^h(a-a_0) + \mathrm{F}_b^h(b-b_0) + \ldots + \mathrm{F}_p^h(p-p_0) = -\mathrm{F}_0^h, \quad h = m+1, m+2, \ldots, m+l. \quad (4)$$

ϕ is now to be minimized with respect to the residuals and corrections, which are $qn+p$ in number and which are subject to $m+l$ conditions; ϕ could therefore be expressed in terms of $qn+p-m-l$ independent variables, and with respect to these it would be an ordinary minimum. Equating to zero the first derivatives of ϕ with respect to these independent variables would give their critical values in the usual way; the critical values of the remaining ones would come from them through (3) and (4). This is the plan outlined by Dr. Stewart*. We shall, however, use Lagrange's scheme of undetermined multipliers. We set the differential of ϕ equal to zero, differentiate (3) and (4), and introduce the multipliers λ_h ($h = 1, 2, \ldots, m, m+1, \ldots, m+l$) in the usual way, and obtain

$$w_i(r_i - x_i) + \sum_{h=1}^{m} \lambda_h \mathrm{F}_i^h = 0, \quad i = 1, 2, \ldots, n, \quad . \quad (5a)$$

$$\sum_{h=1}^{m+l} \lambda_h \mathrm{F}_a^h = 0, \quad \sum_{h=1}^{m+l} \lambda_h \mathrm{F}_b^h = 0, \ldots, \quad \sum_{h=1}^{m+l} \lambda_h \mathrm{F}_p^h = 0, \quad (5b)$$

which with the $m+l$ equations (3) and (4) are sufficient to

* *Loc. cit.*

determine the $qn+p+m+l$ unknowns r_1-x_1, r_2-x_2, ..., $r_{qn}-x_{qn}$, $a-a_0$, $b-b_0$, ..., $p-p_0$, $\lambda_1, \lambda_2, \ldots, \lambda_m, \lambda_{m+1}, \ldots, \lambda_{m+l}$. Equations (6) and (7) are equivalent to (3), (4), (5 *a*), (5 *b*), and constitute the solution. L_{ij} is written for

$$\sum_{k=0}^{q-1} F^i{}_{kn+h}\, F^j{}_{kn+h}/w_{kn+h}. \quad L_{ij} = L_{ji}.$$

The coefficients only are tabled. The unknown for each column stands at the top of that column ; thus the first row is written for the equation

$$L_{11}\lambda_1 + L_{12}\lambda_2 + \ldots + L_{1m}\lambda_m - F_a{}^1(a-a_0) - F_b{}^1(b-b_0)$$
$$- \ldots - F_p{}^1(p-p_0) + 0\,.\,\lambda_{m+1} + 0\,.\,\lambda_{m+2} + \ldots + 0\,.\,\lambda_{m+l} = -F_0{}^1.$$

$$r_i - x_i = -\frac{1}{w_i}\sum_{h=1}^{m} \lambda_h F_i{}^h, \quad i = 1, 2, \ldots, qn. \quad . \quad . \quad (6)$$

λ_1	λ_2	...	λ_m	$a-a_0$	$b-b_0$	...	$p-p_0$	λ_{m+1}	λ_{m+2}	...	λ_{m+l}	
L_{11}	L_{12}	...	L_{1m}	$-F_a{}^1$	$-F_b{}^1$	...	$-F_p{}^1$	0	0	...	0	$F_0{}^1$
L_{21}	L_{22}		L_{2m}	$-F_a{}^2$	$-F_b{}^2$		$-F_p{}^2$	0	0		0	$F_0{}^2$
⋮												⋮
L_{m1}	L_{m2}		L_{mm}	$-F_a{}^m$	$-F_b{}^m$		$-F_p{}^m$	0	0		0	$F_0{}^m$
$F_a{}^1$	$-F_a{}^2$		$-F_a{}^m$	0	0		0	$-F_a{}^{m+1}$	$-F_a{}^{m+2}$		$-F_a{}^{m+l}$	0
$F_b{}^1$	$-F_b{}^2$		$-F_b{}^m$	0	0		0	$F_b{}^{m+1}$	$-F_b{}^{m+2}$		$-F_b{}^{m+l}$	0
⋮												⋮
$F_p{}^1$	$-F_p{}^2$		$-F_p{}^m$	0	0		0	$-F_p{}^{m+1}$	$-F_p{}^{m+2}$		$-F_p{}^{m+l}$	0
0	0		0	$-F_a{}^{m+1}$	$-F_b{}^{m+1}$		$-F_p{}^{m+1}$	0	0		0	$F_0{}^{m+1}$
0	0		0	$-F_a{}^{m+2}$	$-F_b{}^{m+2}$		$-F_p{}^{m+2}$	0	0		0	$F_0{}^{m+2}$
⋮												⋮
0	0	. .	0	$-F_a{}^{m+l}$	$-F_b{}^{m+l}$	...	$-F_p{}^{m+l}$	0	0	...	0	$F_0{}^{m+l}$

. . . . . (7)

These are symmetrical, and it seems appropriate to call them the "general normal eqnations." Since they are linear, the solution is unique. For this statement to hold, it is necessary that the derivatives exist and be single valued ; but such a hypothesis, together with any others essential for complete mathematical rigour, is always fulfilled by the functions that would enter an adjustment by least squares or by any other method. We are now ready to consider applications of the general solution.

For the case of observations on quantities that are subject to m conditions not involving parameters, the normal equations are

$$\left.\begin{aligned} L_{11}\lambda_1 + L_{12}\lambda_2 + \ldots + L_{1m}\lambda_m &= F_0^1 \\ L_{21}\lambda_1 + L_{22}\lambda_2 + \ldots + L_{2m}\lambda_m &= F_0^2 \\ \vdots \qquad\qquad & \vdots \\ L_{m1}\lambda_1 + L_{m2}\lambda_2 + \ldots + L_{mm}\lambda_m &= F_0^m \end{aligned}\right\} \quad . \quad . \quad (8)$$

The corrections, $r_i - x_i$, are, as before, given by (6), and can be computed as soon as $\lambda_1, \lambda_2, \ldots, \lambda_m$ have been obtained from (8). The equations (8) with (6) were known to Gauss. They are usually called the "solution by correlates." In problems that come under this case, usually only one nature of quantity enters, in which event $q = 1$.

In order to use the general solution for curve fitting, we let the conditions (1) or (3) express the fact that the most probable values of observed quantities and of parameters must satisfy the empirical formula exactly. Suppose the formula is $F(x, y, z, \ldots ; a, b, \ldots, p) = 0$; then, if the qn observations were wholly devoid of error, they would satisfy

$$F(x_h, x_{n+h}, \ldots, x_{(q-1)n+h} ; a, b, \ldots, p) = 0, \quad h = 1, 2, \ldots, n. \quad (9)$$

m now becomes equal to n. If we adopt the notation

$$F^h(x_1, x_2, \ldots, x_{qn} ; a, b, \ldots, p)$$
$$= F(x_h, x_{n+h}, \ldots, x_{(q-1)n+h} ; a, b, \ldots p),$$

then the conditions (1) or (3) become the requirement that the most probable values satisfy the formula exactly. Clearly, then,

$$F_g^h = F^h_{n+g} = F^h_{2n+g} = \ldots = 0 \quad \text{and} \quad L_{gh} = 0 \text{ if } g \neq h.$$

If the empirical curve (or surface) is known to pass rigorously through l given points, this fact can be expressed by (2) or (4). If there are no relations between the parameters other than (1), then (2) and (4) and the multipliers $\lambda_{m+1}, \ldots, \lambda_{m+l}$ do not exist. The general solution now simplifies to the following equations :—

$$r_{kn+h} - x_{kn+h} = -\frac{1}{w_{kn+h}} \lambda_h F^h_{kn+h}, \quad \left\{\begin{matrix} k = 0, 1, \ldots, q-1 \\ h = 1, 2, \ldots, n \\ kn+h = 1, 2, \ldots, qn \end{matrix}\right\}.$$
$$. \quad . \quad . \quad . \quad (10)$$

$$L_h = \sum_{k=0}^{q-1} (F^h{}_{kn+h})^2 / w_{kn+h}, \qquad \left\{ \begin{matrix} k = 0, 1, \ldots, q-1 \\ h = 1, 2, \ldots, n \\ kn+h = 1, 2, \ldots, qn \end{matrix} \right\}. \quad (11)$$

$$\begin{array}{cccccccccc}
\lambda_1 & \lambda_2 & \ldots & \lambda_n & a-a_0 & b-b_0 & \ldots & p-p_0 & \lambda_{n+1} & \lambda_{n+2} & \\
L_1 & 0 & \ldots & 0 & -F_a{}^1 & -F_b{}^1 & \ldots & -F_p{}^1 & 0 & 0 & F_0{}^1 \\
0 & L_2 & & 0 & -F_a{}^2 & -F_b{}^2 & & -F_p{}^2 & 0 & 0 & F_0{}^2 \\
\vdots & & & & & & & & & & \vdots \\
0 & 0 & \ldots & L_n & -F_a{}^n & -F_b{}^n & \ldots & -F_p{}^n & 0 & 0 & F_0{}^n
\end{array} \quad \ldots \quad (12)$$

$$\begin{array}{ccccccc}
\Sigma \frac{F_a{}^h F_a{}^h}{L_h} & \Sigma \frac{F_a{}^h F_b{}^h}{L_h} & \ldots & \Sigma \frac{F_a{}^h F_p{}^h}{L_h} & F_a{}^{n+1} & F_a{}^{n+2} & -\Sigma \frac{F_a{}^h F_0{}^h}{L_h} \\
\Sigma \frac{F_b{}^h F_a{}^h}{L_h} & \Sigma \frac{F_b{}^h F_b{}^h}{L_h} & & \Sigma \frac{F_b{}^h F_p{}^h}{L_h} & F_b{}^{n+1} & F_b{}^{n+2} & -\Sigma \frac{F_a{}^h F_0{}^h}{L_h} \\
\vdots & & & & & & \vdots \\
\Sigma \frac{F_p{}^h F_a{}^h}{L_h} & \Sigma \frac{F_p{}^h F_b{}^h}{L_h} & & \Sigma \frac{F_p{}^h F_p{}^h}{L_h} & F_p{}^{n+1} & F_p{}^{n+2} & -\Sigma \frac{F_a{}^h F_0{}^h}{L_h} \\
F_a{}^{n+1} & F_b{}^{n+1} & & F_p{}^{n+1} & 0 & 0 & -F_0{}^{n+1} \\
F_a{}^{n+2} & F_b{}^{n+2} & & F_p{}^{n+2} & 0 & 0 & -F_0{}^{n+2}
\end{array} \quad \ldots \quad (13)$$

Σ denotes summation with h running from 1 to n.

In practice, the curve is seldom required to pass rigorously through more than one point, but the equations have been written with $l = 2$ for generality. They are easily extended if $l > 2$. If $l = 0$, the $(p+1)$th and $(p+2)$th columns and rows are to be deleted. Again we have symmetrical equations; in fact the symmetry extends through (12) and (13) together. They are given separate numbers because $\lambda_1, \lambda_2, \ldots, \lambda_n$ do not occur at all in equations (13). The computation can be stopped as soon as (13) have been solved if the corrections $r_i - x_i$ are not desired. Since (13) involve the parameters only, the name "p normal equations" is suggested for them. (12) and (13) together could be called the "normal equations."

In all this work it should be noted that x_{kn+h} with $k = 0, 1, 2$ becomes x_h, x_{n+h}, x_{2n+h} or, if preferred, x_h, y_h, z_h. The number of terms in each L_h cannot be greater than q,

which is seldom greater than 3 ; and if any weight w_{kn+k} is infinite, the number of terms in L is further reduced.

Let us see what happens when we make the following specifications :—(a) $x_1, x_2, \ldots, x_n$ subject to error, but $x_{n+1}, \ldots, x_{qn}$ relatively free from error. This means that the weights $w_{n+1}, \ldots, w_{qn}$ are infinite compared with $w_1, \ldots, w_n$, which makes the corrections $r_{n+1} - x_{n+1}, \ldots, r_{qn} - x_{qn}$ zero and throws all the adjustment on to $x_1, \ldots, x_n$. (b) The empirical formula is solved explicitly for x_k $(k = 1, 2, \ldots, n)$, so $F_k^k = 1$ and $L_k = 1/w_k$. (c) $l = 0$. Equations (13) then become

$$\left.\begin{aligned}
\Sigma F_a{}^k F_a{}^k w_k (a - a_0) + \Sigma F_a{}^k F_b{}^k w_k (b - b_0) + \ldots + \Sigma F_a{}^k F_p{}^k w_k (p - p_0) \\
= -\Sigma F_a{}^k F_0{}^k w_k \\
\Sigma F_b{}^k F_a{}^k w_k (a - a_0) + \Sigma F_b{}^k F_b{}^k w_k (b - b_0) + \ldots + \Sigma F_b{}^k F_p{}^k w_k (p - p_0) \\
= -\Sigma F_b{}^k F_0{}^k w_k \\
\vdots \qquad\qquad\qquad\qquad \vdots \\
\Sigma F_p{}^k F_a{}^k w_k (a - a_0) + \Sigma F_p{}^k F_b{}^k w_k (b - b_0) + \ldots + \Sigma F_p{}^k F_p{}^k w_k (p - p_0) \\
= -\Sigma F_p{}^k F_0{}^k w_k
\end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (14)$$

Equations (14) will be found in most textbooks, where they are expected to take care of any problem in curve fitting. The texts do not generally state the three stipulations that equations (14) are subject to. It will be recognized that w_k is often called the " weight of the k-th observation equation," a phrase to which no meaning can be ascribed unless only one quantity in an " observation equation " is subject to error. The two phrases in quotation marks have caused much confusion. Most texts do not warn the reader that problems may arise for which equations (14) are not applicable ; much less do they tell how to handle such a situation. In fact some books do not mention weights at all ! and most of them are full of problems upon which a solution by least squares cannot even be commenced because there is no information given concerning the weights of the observations to be adjusted. The mumpsimus of throwing the adjustment on to the wrong set of observations has many times caused least squares to be accused of yielding unreasonable results. The general equations (12) and (13) must be followed if an adjustment by least squares is desired. The normal equations (12) and (13) are affected by various assignments of weights through the L_k.

Suppose we have $2n$ observations x_h, y_h (or x_h, x_{n+h}) of weights w_h, w_{n+h} ($h = 1, 2, \ldots, n$), which, if devoid of error, would satisfy the empirical formula $y = ax + b$, which must pass through the point x_0, y_0. We write

$$F^h(x_1, \ldots, x_{2n}\ ;\ a, b) \equiv y_h - ax_h - b, \quad h = 1, 2, \ldots, n;$$
$$F^{n+1}(a, b) \equiv y_0 - ax_0 - b.$$

After the approximations a_0 and b_0 have been determined, we have

$$F_h{}^h = -a_0, \quad F^h{}_{n+h} = 1, \quad F_a{}^h = -x_h, \quad F_b{}^h = -1,$$
$$F_0{}^h = y_h - a_0x_h - b_0, \qquad F_0{}^{n+1} = y_0 - a_0x_0 - b_0,$$
$$F_a{}^{n+1} = -x_0, \quad F_b{}^{n+1} = -1, \quad L_h = a_0{}^2/w_h + 1/w_{n+h}.$$

Equations (13) now give

$$\left.\begin{aligned} \Sigma \frac{x_h{}^2}{L_h}(a-a_0) - \Sigma \frac{x_h}{L_h}(b-b_0) + x_0\lambda_{n+1} &= \Sigma \frac{x_h F_0{}^h}{L_h} \\ \Sigma \frac{x_h}{L_h}(a-a_0) + \Sigma \frac{1}{L_h}(b-b_0) + \lambda_{n+1} &= \Sigma \frac{F_0{}^h}{L_h} \\ x_0(a-a_0) + \qquad (b-b_0) + \qquad 0 &= -F_0{}^{n+1} \end{aligned}\right\}. \qquad (15)$$

These are the "p normal equations" for fitting a straight line that must pass rigorously through the point x_0, y_0. If this condition is removed, the third column and third row disappear. The adjusted values of the observations can easily be obtained by computing λ_1, λ_2, ..., λ_n from (12), and thence $r_i - x_i$ from (10); $x_i + (r_i - x^i)$ gives the adjusted value (most probable value) of the quantity that was observed to be x_i.

An empirical formula defines a curve or surface. x_h, x_{n+h}, x_{2n+h}, ... are the observed coordinates of a point whose most probable position is at r_h, r_{n+h}, r_{2n+h}, .. , which lies on the surface. Referred to q orthogonal axes 01, 02, ..., 0Q, the direction cosines of the line H joining the observed and most probable positions of this point are in the ratio $r_h - x_h : r_{n+h} - x_{n+h} : r_{2n+h} - x_{2n+h} : \ldots$, which, on account of (10), are the same as $F_h^h/w_h : F_{n+h}^h/w_{n+h} : F_{2n+h}^h/w_{2n+h} : \ldots$. But $F_h^h : F_{n+h}^h : F_{2n+h}^h : \ldots$ are (except for powers of the residuals) the ratios of the direction cosines of the normal to the surface at the point r_h, r_{n+h}, r_{2n+h}, ... So if $w_h = w_{n+h} = w_{2n+h} = \ldots$, the line H is normal to the surface. If $w_h/w_{n+h} = w_h/w_{2n+h} = \ldots = 0$, all the adjustment is thrown on to x_h and none on to x_{n+h}, x_{2n+h}, ..., and the line H is parallel to the 01 axis. These features for two dimensions

were recently noticed by the writer * in considering the fitting of a series of points by a straight line, and they are now seen to be general properties.

The following problem illustrates the various results one gets by assigning various weights to the observations. The observed x and y coordinates of five points, supposed to lie on a circumference, are $(-5{\cdot}80, 9{\cdot}10)$, $(-1{\cdot}60, 9{\cdot}82)$, $(3{\cdot}20, 8{\cdot}16)$, $(5{\cdot}63, 3{\cdot}80)$, $(5{\cdot}87, -{\cdot}62)$. Determine their most probable circle (i.) when the x and y coordinates all have equal weight; (ii.) when the x coordinates are free from error and the y coordinates all have equal weight; (iii.) when the y coordinates are free from error and the x coordinates all have equal weight. Let

$$(x-a)^2+(y-b)^2-c^2=0$$

be the equation that the observed coordinates would satisfy if they were free from error. a, b, c are sought. Approximate values $a_0=-2{\cdot}25$, $b_0=1{\cdot}50$, $c_0=8{\cdot}3$ were estimated by graph.

We write x_{5+h} for y_h and

$$F^h(x_1, x_2, \ldots, x_{10};\ a, b, c) \equiv (x_h-a)^2+(x_{5+h}-b)^2-c^2.$$
$$F_h^h = 2(x_h-a_0), \qquad F_{5+h}^h = 2(x_{5+h}-b_0),$$
$$F_a^h = -2(x_h-a_0), \quad F_b^h = -2(x_{5+h}-b_0), \quad F_c^h = -2c_0.$$

When weights are equal they may all be put equal to unity, so

$$L_h = 4(x_h-a_0)^2+4(y_h-b_0)^2 \qquad \text{for case (i.)};$$
$$L_h = 4(y_h-b_0)^2 \qquad \text{for case (ii.)};$$
$$L_h = 4(x_h-a_0)^2 \qquad \text{for case (iii.)}.$$

The p normal equations (13) were set up for the three cases. The results, kindly calculated by Miss Lola E. Shupe, are tabulated below.

Case.	$a-a_0$.	$b-b_0$.	$c-c_0$.	a.	b.	c.
(i.).....	·070	·135	−·020	−2·179	1·635	8·280
(ii.).....	−·232	−·180	·246	−2·482	1·320	8·546
(iii.).....	·034	−·008	·055	−2·216	1·492	8·355

One need not draw these to see that they are three distinct circles; a glance at the different values obtained for a, b, c is sufficient.

In case (i.) the most probable position (r_h, r_{5+h}) of a point observed to be at (x_h, x_{5+h}) is where a line joining (x_h, x_{5+h})

* W. Edwards Deming, Phys. Soc. Proc. xlii. p. 97 (1930).

with the centre $(-2{\cdot}179, 1{\cdot}635)$ cuts the circle determined by (i.) ; this is in accordance with the requirement found above that when the weights of x_h, x_{n+h}, x_{2n+h}, ... are equal, the line joining the observed and most probable positions of a point is normal to the graph of the formula. In case (ii.) the most probable position is where a line through (x_h, x_{5+h}) parallel to the y axis cuts the second circle ; and in (iii.) it is where a line through (x_h, x_{5+h}) parallel to the x axis cuts the third circle. The three different assignments of weights thus give three distinct circles and different locations for the most probable positions of the five points. This example illustrates the importance of the weights of the observations. It also illustrates how least squares, being the development of a theory of errors, is able to take account of weights, or rather, depends on weights for its application.

The foregoing material is a consequence of the statement of the principle of least squares. The aim has been to show the principle is applied to different types of problems. There has been no desire to provoke any argument concerning the validity of the principle nor the practicability of its application. Before its results can be discussed, it must be properly applied, and this paper is an attempt to assist in its application. The assignment of weights appears to be fundamental. Once the weights are assigned, the procedure is quite definitely marked out, and the results are unambiguous. The question of just what values should be assigned to the weights in the first place is often not so clear cut. I venture to make a few remarks on this phase of the subject.

For n observations all of the same nature, it may not be difficult for the observer to decide what weight should be given to each observation, and though it may be largely up to his judgment at times, still it can be accomplished satisfactorily. The assignment of weights is a powerful tool, and experience is necessary in order to use it properly. Weights are relative, and when there is only one nature of quantity that is subject to error, the effect of doubling every weight is merely to multiply every p normal equation through by two, which leaves them unaltered. But when the L_h in (13) consist of more than one term, the situation is more complicated. If we double each of the w_h $(h=1, 2, \ldots, n)$ and leave $w_{n+h}, w_{2n+h}, \ldots$ fixed, the p normal equations are altered. So it is not only necessary that the n weights in the set $w_1, w_2, \ldots, w_n$ bear the proper relation to each other, and likewise for the set of n weights $w_{n+1}, \ldots, w_{2n}$, etc., but there must be some relation

between the sets. In other words, only one of the qn weights can be fixed arbitrarily; the others bear some relation to it. Thus, if we have measurements on the pressure, volume, and temperature of a given mass of gas at n different temperatures, not only must the weights of the observations P_1, P_2, ..., P_n (and those of V_1, V_2, ..., V_n, and those of T_1, T_2, ..., T_n) bear the proper ratio to one another, but the weights of P_h, V_h, T_h must also have some relation to each other at $h=1, 2, \ldots, n$. One way of settling upon the proper values for the weights of observations of different natures is to keep in mind the geometry of curve fitting, remembering that if the weights w_h, w_{n+h}, w_{2n+h} are equal, the line joining the observed and most probable positions of the point is normal to the empirical surface. If one has some idea concerning how the line joining the observed and most probable positions of a point should be drawn, one can decide how the weights should be assigned. The adjustment of a set of observations and the determination of parameters, by least squares, cannot be commenced until each observation is allotted a weight. If the task of assigning weights to the observations is for some reason impossible, then some method other than least squares must be employed.

XII. *Some Further Experiments on the Combustion of Inflammable Gases by Electric Sparks. By* J. D. Morgan, *D.Sc.**

IT has previously been shown† that the ignition of an inflammable gas mixture by an electric spark is consistent with the hypothesis that ignition depends on the heating of a sufficient volume of the gas by conduction to a sufficient temperature.

In the previous experiments each ignition was effected by a single spark, and the spreading of the flame through the gas after the spark had done its work was due to the burning of the gas itself. The experiments were carried out on mixtures lying within the so-called limits of inflammability, that is to say, mixtures capable of sustaining the flame initiated by the spark. Mixtures lying beyond

* Communicated by the Author.

† Taylor Jones, Morgan, and Wheeler, Phil. Mag. xliii. p. 359 (February 1922).

the upper and lower limits of inflammability are capable of being ignited by a spark, but in such mixtures the flame dies out before the whole of the gas is consumed.

It has long been known that any spark, however small, causes some combustion of the surrounding gas, whether the mixture lies within or outside the limits of inflammability, though complete inflammation of mixtures lying within those limits only occurs when the spark is sufficiently energetic. Sometimes a spark which is insufficient to produce inflammation of the whole of the gas is seen to be surrounded by a small flame, but it is quite easy to arrange the conditions of the experiment so that no flame is visible around the spark and yet combustion occurs. In this case, by continuing the spark-discharge for a sufficient time, nearly the whole of the gas can be consumed without any appearance of flame.

As an extension of the earlier experiments, it was thought desirable to ascertain whether the flameless combustion of inflammable gas-mixtures by spark-discharges was also consistent with the hypothesis above mentioned, and in the following a short account is given of some of the results obtained.

The procedure adopted was to use a very weak mixture of coal-gas and air through which a rapid succession of high-tension sparks could be maintained without any appearance of burning in the gas, the latter being at atmospheric pressure and temperature. Within a small combustion chamber was arranged a variable spark-gap, and this was connected to the secondary winding of an induction coil, the primary of which was controlled by a rapidly moving interrupter of the trembler type. Various forms of spark-gap electrodes were tried, and as it was thought to be desirable to keep as small as possible the metallic area exposed to the gas, a needle-point was used as one electrode and a rod of about 0·125 inch diameter having a hemispherical end, as the other. The needle-point was connected to the negative terminal of the induction coil and the rod to the other terminal, which was earthed. It was found that the sparking voltage was more regular with this combination of electrodes than in a gap formed by a pair of needle-points. Another feature of the apparatus was a resistance of about 1500 ohms placed in series with the earthed electrode. As is well known, the spark-discharge from an induction coil

is of complex character, consisting of an initial bright discharge, termed the capacity component, and a subsequent flame, termed the inductance component. It has been found that the capacity component is more effective as an igniting agent than the inductance component, and as it was felt to be desirable to reduce the capacity component as much as possible, so that a rapid succession of sparks sustained for a given interval of time would approximate to a uniform electrical flame or arc, the resistance above mentioned was inserted for the purpose of minimizing the capacity component. The heating effect produced by the spark-discharge in the combusion chamber when filled with air was measured by the pressure indicated on a water-pressure gauge. The same gauge was also used for measuring the amount of combustion which had taken place after the discharge had been passed through the combustible gas-mixture. As combustion of a coal-gas and air-mixture results in a diminution of volume, the reduction of the pressure in the chamber at the completion of the experiment afforded a convenient measure of the amount of combustion that had taken place.

There are obviously two ways in which the conditions associated with the spark-discharge can be varied. On the one hand the gap can be kept constant, and the amount of heat produced in the discharge varied by varying the current in the primary winding of the induction coil. On the other hand, the primary current can be kept constant and the gap-width varied. It was found that over a considerable range the heating effect produced by the discharge remains constant while the gap width is being increased (the primary current being constant), but when the gap is increased beyond a certain limit the heating effect falls, and continues to fall until the width of the gap is such that no spark can pass.

In the experiments the same gas mixture was used throughout, namely, a very weak coal-gas and air-mixture through which the spark-discharge could be sustained without the appearance of flame, and in each experiment the discharge was sustained for one minute. After each experiment the chamber was evacuated and another charge was introduced.

The results obtained with a constant gap-width and varying the primary current were such as might be

anticipated from previous experience. The amount of gas consumed was proportional to the heating effect of the discharge. The following is a typical set of observations :—

Heating effect *.	Gas consumed *.
2·25	3
5·5	7·5
8	10·5
14	19

* The figures are divisions on the pressure-gauge scale, each division being 0·125 inch.

The results obtained on keeping the primary-current constant and varying the gap-width were not so easy to understand. It was found that the amount of gas consumed was proportional to the gap-width so long as the heating effect of the discharge remained constant. When the heating effect fell, in consequence of excessive increase of gap-width, the gas consumption also fell. The following is a typical set of observations :—

Gap-width.	Gas consumed.
1 unit *	4·5
2 units	9·0
3 ,,	13·5
4 ,,	18·0

* 1 unit=0·05 inch.

In an earlier paper † it was shown that the spark-energy required to ignite a given inflammable gas-mixture by means of a single spark varied largely with the gap-width, and it was also shown that this result was consistent with the thermal hypothesis already mentioned. The relationship found over a considerable range between the spark-energy (j) required to produce ignition and the gap-width (x) can be expressed approximately by $j.x=c$, where c is a constant. It seems reasonable to suppose that, as the spark in each case was the least that would cause ignition under the conditions of the experiment, the effect produced in the gas was always the same, notwithstanding that variation in spark-length necessitated a change in the amount of spark-energy required to produce that effect. On the thermal hypothesis

† J. D. Morgan, Phil. Mag. xlv. p. 968 (May 1923).

ignition depends upon raising the outer boundary of a definite volume of the gas to a definite temperature. In the theoretical development of this hypothesis, already referred to, it was assumed that the necessary boundary temperature was the ignition temperature of the gas. The hypothesis is therefore equivalent to supposing that ignition depends upon the combustion by the spark of some definite quantity of the gas. On the assumption that this quantity, for a given gas-mixture, is always the same, the expression $j \cdot x = c$ can be written $j \cdot x = a \cdot q$, where a is a constant and q the quantity of gas consumed by the spark. On this basis the results found in the experiments now being described are readily intelligible. Keeping x constant and varying j, the quantity of gas consumed should be proportional to j, whilst when j is kept constant and x is varied, the gas consumed should be proportional to x, and this is what was found. It may be concluded, therefore, that the combustion of a gas-mixture by electric sparks is consistent with the hypothesis that the combustion of the gas depends on the heating effect of the sparks.

There appears to be no justification for assuming or supposing that a spark behaves towards the gas in any way differently from that of any other hot body. Using the same gas-mixture as in the spark experiments, the spark was replaced by a length of wire heated by an electric current. A current of 1·0 amp., maintained for five seconds, resulted in a gas-consumption corresponding to 16 divisions on the pressure-gauge scale, and a current of 1·25 amp. resulted in a gas-consumption represented by 25 divisions. It will be seen that the gas-consumption was proportional to the square of the current. The heating effect of the hot wire was also proportional to the square of the current. As the amount of gas consumed by a spark in a gap of constant width and by a heated wire was proportional to the heat produced, it must be inferred that the action on the gas of both the hot wire and the spark was the same.

Conclusion.

The experiments described show that when a succession of electric sparks is maintained in a gas-mixture which cannot be ignited or inflamed by the sparks, the amount of combustion, or combination of the constituents,

which takes place in the mixture is proportional to the heat-energy of the spark-discharge when the gap-width is kept constant, and is proportional over a wide range to the gap-width when the heat-energy of the discharge is kept constant, and the deduction made from these experiments, on the basis of earlier experiments, is that the results are consistent with the hypothesis that combustion, like ignition, depends upon the heating of the gas by conduction.

XIII. *High Frequency and Direct Current Discharges in Helium. By* F. Llewellyn Jones, *B.A., Demonstrator at the Electrical Laboratory, Oxford* *.

1. In the course of the experiments forming a general investigation of the electrical properties of the monatomic gases that have been made in the Electrical Laboratory, Oxford, it was found that in many cases impurities (which may be present in quantities too small to be detected spectroscopically) have a definite effect on the conductivity.

In tubes fitted with internal electrodes small quantities of impurity are evolved from the metal parts under the action of the discharge. This effect is avoided by using high frequency discharges in tubes with external electrodes in the form of sleeves which can slide over the tube, and whenever possible it is of advantage to use this method of producing the discharge.

2. It has been shown by Townsend and Donaldson † that the potentials V between the electrodes required to maintain various currents in high frequency discharges increase with the current, but for a given current the rate of increase of the potential with the distance x between the electrodes is a constant. The constant $\frac{dV}{dx}$, which represents the force F in the uniform glow, is the same over a large range of currents and over a large range of frequencies of oscillation but it depends on the pressure of the gas.

* Communicated by Prof. J. S. Townsend, F.R.S.

† Phil. Mag. v. p. 178 (Jan. 1928).

This result, which was obtained experimentally with nitrogen, helium, and neon, formed the basis of a theory of the uniform glow in the electrodeless discharge *. This theory indicates that the mean value of the electric force in the uniform glow of the high frequency discharge should be the same as the force in the uniform positive column of a continuous current discharge over a large range of pressures.

The experiments described in this paper confirm the theory in showing that the mean force in the uniform glow in high frequency discharges in helium is the same as in the uniform positive column of direct current discharges, for pressures from 1 millimetre to 35 millimetres, in a tube 2·9 centimetres internal diameter.

3. The range of oscillation frequencies within which the correspondence between the high frequency discharge and the uniform positive column may be expected to hold is determined, approximately, from a consideration of the mean energy of agitation E of the electrons in the gas.

The mean energy of agitation of the electrons in the positive column of a direct current discharge in a wide tube may be obtained from measurements of the electric force X and the gas pressure p, since the values of E have been determined in terms of the ratio X/p. The value of E in volts is $k/27$, and the values of k have been found in terms of the ratio X/p for several gases †.

In order that the same light and the same rate of ionization may be obtained in the uniform luminous column of a high frequency discharge the mean energy of agitation of the electrons must be the same as in the direct current discharge. This result is obtained when the mean force in the high frequency oscillation is equal to the force X in the direct current discharge and the period T of the high frequency oscillation is within a certain range. This range is very wide, and the lower and upper limits of the wave-lengths of the oscillation may be found from the following considerations.

In order that the energy of agitation of the electrons moving under the action of an oscillating force, $F = A \sin 2\pi t/T$, may be the same as the energy attained under a direct force X, where the mean value of $F\ (2A/\pi)$ is equal

* J. S. Townsend, *Comptes Rendus*, t. clxxxvi. p. 55 (Jan. 1928).

† J. S. Townsend, 'Motions of Electrons in Gases' (Clarendon Press).

to X, the number of free paths of the electrons in which the force F changes in direction must be small compared with the number of free paths traversed by the electrons in the time T/2. This condition is satisfied when T/2 is large compared with the time required by an electron to traverse the mean free path. Thus, T must be large compared with $2l/U$, where U is the mean velocity of agitation of the electrons and l is the mean free path. If U be taken as $1{\cdot}3 \times 10^8$ (which is the velocity of agitation of the electrons when their kinetic energy corresponds to 4 volts *), and the mean free path of the electrons in helium at 1 millimetre pressure be taken as ·05 centimetre†, then $2l/U$ is 8×10^{-10} second, and the above condition is realised when T exceeds 4×10^{-8} second, which is the period of wave-length of 12 metres. This may be taken as the order of the lower limit to the range of wave-lengths under consideration, when the pressure of the gas is about 1 millimetre. The above calculation shows that this limit is inversely proportional to the pressure.

Several effects have to be taken into consideration in order to assign an upper limit to the range. For example, the distance traversed by the electrons in an ordinary tube in the direction of the axis in a time T/2 should not exceed 1 or 2 centimetres. Since the mean velocity W in the direction of the electric force is about U/100, it follows that TU must be less than 200, or T must be less than $2/1{\cdot}3 \times 10^{-6}$, so that the upper limit of the range may be about 600 metres.

The electric force in the uniform glow has been determined experimentally over a large range of wave-lengths by Hayman ‡, and he found that the force in discharges in helium and neon was the same for wave-lengths from 40 to 640 metres. The gas pressures in his experiments were from 1 to 20 millimetres in helium and from 1 to 80 millimetres in neon.

4. The helium used in the experiments which are described in this paper was prepared by heating thorianite contained in a quartz tube. The gas was dried and passed into tubes containing charcoal, which were cooled by liquid air, and after remaining for some hours in these

* The energy of the electrons in the positive column in helium is of this order. See J. S. Townsend, Phil. Mag. ix. p. 1145 (June 1930).

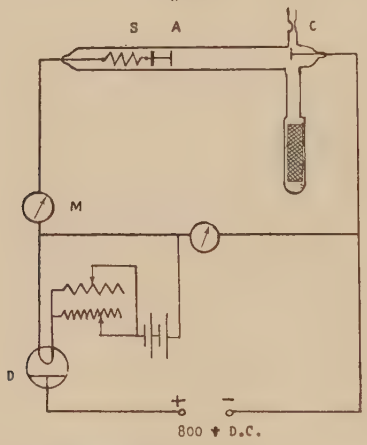
† 'Motions of Electrons in Gases.'

‡ Phil. Mag. vii. p. 586 (March 1929).

tubes most of the impurities were removed and the gas was pumped into a flask. These operations were repeated until a sufficient quantity was collected for the experiments.

The method of eliminating small traces of impurity from the gas before admitting it into the discharge-tube was the same as that used in other researches carried out in the laboratory. Traces of hydrogen were removed by passing the gas, which was taken from the flask, over copper oxide kept at a dull red heat. Other impurities were removed by keeping the gas in a tube containing charcoal and cooled by liquid air for sixteen hours before

Fig. 1.



admitting it to the discharge-tube. The details of the precautions taken in these operations have already been explained in papers to which reference must be made *.

The purity of the gas was first tested by a spectroscopic examination of the high frequency discharge, and then by measurements of the force in the uniform positive column of the direct current discharge.

5. The discharge-tube is shown in fig. 1. It was made of pyrex glass, and the electrodes were made of aluminium. The tube was 30 centimetres long and 2·9 centimetres internal diameter.

The cathode was a disk 2·7 centimetres in diameter, and

* J. S. Townsend and C. M. Focken, Phil. Mag. ii. p. 474 (1926); L. G. H. Huxley, Phil. Mag. v. p. 721 (April 1928).

was fixed to a piece of tungsten wire which was brought through the glass. The anode A was attached to a light iron magnet and was connected to a tungsten seal by means of a length of thin aluminium strip S. This enabled the distance between the electrodes to be varied by sliding the anode inside the tube by means of a magnet. Near the cathode was fitted a side tube (about 2 centimetres in diameter and 15 centimetres long) containing charcoal enclosed in copper gauze. During the experiments the charcoal was cooled by liquid air in order to absorb impurities given off by the electrodes while a direct current was passing.

The tube was thoroughly baked out before use, and, together, with the charcoal, it was frequently heated in order to remove impurities which accumulate during the course of several experiments. This was done with helium in the tube in order more easily to heat the electrodes and to carry away impurities as the tube was being evacuated. Evacuation was always continued until the tube and charcoal had cooled throughly. Impurities given off from the electrodes and from the walls were thus removed.

6. The current in the gas was measured by the milliammeter M, and the potential difference between the electrodes was given by the voltmeter V. The source of high potential was an 800-volt direct current generator. The current could be adjusted to any desired value by means of a two-electrode valve D interposed in the positive lead from the generator. The potentials V_1 and V_2 were measured with the electrodes at the distances 14·5 and 17·5 centimetres apart. The experiments were performed throughout the range of pressures from 0·5 to 35 millimetres with different samples of gas and with different currents from 1 to 8 milliamperes.

The electric force $(V_2—V_1)/3$ was found to be independent of the current for pressures greater than 2 millimetres and to be independent of the length of the uniform positive column, but it depended on the pressure of the gas and was affected by the presence of impurities.

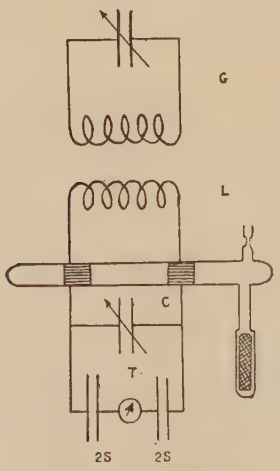
The results of the experiments with pure helium are shown in the following table, where the pressures p of the gas in millimetres are given in the first column and the electric force X in the second column.

7. The electrical arrangements for measuring the potentials in the high frequency discharge were similar to those developed by Townsend and Donaldson * and used by Hayman † in his experiments with helium and neon.

Helium.

p.	X.	$2A/\pi$.	X/p.	E.
35	43·5	41·0	1·2	2·4
22	33·3	32·4	1·5	3·0
17	27·2	27·9	1·6	3·2
13	21·0	22·4	1·7	3·4
10	17·6	18·5	1·85	3·7
6	11·0	10·7	1·9	3·8
3·8	8·0	7·7	2·1	4·1
2·6	6·6	6·0	2·5	4·5
2·0	6·0	5·8	3·0	5·2

Fig. 2.



The source of the high frequency oscillations was a 60-watt valve generator G (fig. 2). This was loosely coupled to the secondary circuit (L, C S) which was approximately in tune with the generator.

The external electrodes consisted of lead sleeves 4

* *Loc. cit.* † *Loc. cit.*

centimetres long which could slide over the tube to any desired distance apart, and these were connected to the ends of the coil L.

The amplitude of the potential V between the electrodes was found in terms of the current passing through the two condensers, each of capacity 2S. The current was measured by the thermal ammeter T connected between the two condensers 2S, so that V is given by the formula, $V = I/2\pi fS$, where f is the frequency of the oscillations, and I the amplitude of the current.

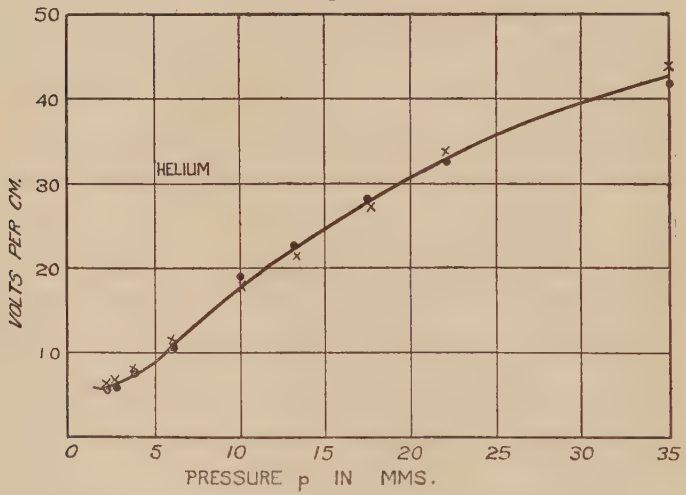
All these experiments were performed with the tube which was used for the direct currents, and the wave-length of the oscillations was 45 metres. As it was found in previous experiments that the force in the uniform high frequency glow was independent of the wave-length, it was considered unnecessary to measure the forces in the high frequency discharge for more than one frequency of oscillation.

The potential difference between the electrodes was measured when the sleeves were at distances of 3 and 6 centimetres apart and the same current passing through the tube. The intensity of the current was indicated by the effect of the light from the discharge on a selenium cell, and the current was adjusted by varying the strength of the oscillations in the generator until the resistance of the selenium cell had the same value when the electrodes were at different distances apart. In the course of the direct current experiments the effect of the light from the discharge on the selenium was noted, so that the current could be indicated by the resistance of the selenium. It was thus possible to adjust the currents in the high frequency discharges to any value from 1 to 8 milliamperes.

The results of the experiments are given in the third column of the table, where the values of the mean force $2A/\pi$ are expressed in volts per centimetre. The agreement between the forces in the uniform glow of the high frequency and direct current discharges can be seen. These results are also shown by the curve in fig. 3, where the ordinates are the force in the uniform glow and the abscissæ the pressure p of the gas. The points marked with circles represent the mean force $2A/\pi$ in the high frequency discharges, and the points marked with crosses the force X in the direct current discharges, as found experimentally.

8. The numbers given in the table show that the ratio X/p diminishes as p increases. The ratios X/p obtained from the curve (fig. 3) are given in the third column of the table, and the values of the mean energy of agitation E of the electrons in volts corresponding to these values of X/p * are given in the fifth column. It is seen that E increases as p diminishes. It will be noticed that at a pressure of 35 millimetres the mean energy of agitation of the electrons is not greater than 2·4 volts, and yet the discharge is maintained by the process of ionization by collision. Also, some collisions of the electrons with atoms of the gas cause the emission of light.

Fig. 3.



As the mean energy of agitation of the electrons increased the colour of the discharge changed. At the higher pressures the red and yellow lines in the spectrum predominated, and at the lower pressures the green and blue lines became intense. The colour of the direct discharge and the distribution of energy in the different lines of the spectrum was substantially the same as in the high frequency discharge at the same pressure.

Thus the experiments show that the greater the energy of agitation of the electrons the greater the intensity of the lines near the violet end in the visible spectrum. (This may also be deduced from the fact that in parts of the

* 'Motions of Electrons in Gases.'

discharge near the electrodes where the electric force was much greater than in the long luminous column the green and blue lines were very intense and those near the red end of the spectrum were very faint.)

The intensity of the light was greater near the axis of the tube than at the surface. Also the colour was different. In the light from points near the axis the relative intensity of the lines near the violet end of the spectrum was greater than in the light from points near the surface. This is in accordance with a theory due to Townsend *. This theory shows that there is a radial force due to a positive charge in the gas which repels positive ions towards the surface and retards electrons in their diffusion to the surface. Thus the mean energy of agitation of the electrons near the axis is greater than that of the electrons near the surface, and the experiments show that the greater the energy of agitation the greater the intensity of the lines near the violet end of the visible spectrum. Also the collisions of the electrons with the walls of the tube tend to lower their mean energy of agitation at the surface.

A more complete investigation of the intensities of the different lines in the spectrum for different pressures of the gas is at present being made, and the results will shortly be published in another paper.

9. Impurities in quantities too small to be easily detected by a direct-vision spectroscope produced a definite effect on the force in the positive column and in the high frequency glow. The following results obtained with a sample of gas at 13 millimetres indicate how the force was affected.

At first, when the gas was pure the force was found to be 21 volts per centimetre. A large direct current was then passed for about ten minutes, which caused small quantities of impurities to be given off. The force was then found to be 18 volts per centimetre. After about fifteen minutes the charcoal had absorbed most of the impurities, and the force was restored to the higher value of 21 volts per centimetre.

The direct current discharge was very effective in sweeping impurities into the space near the cathode, where

* 'Electricity in Gases,' p. 440; also *Comptes Rendus*, t. clxxxvi. p. 55 (Jan. 1928).

they were absorbed by the charcoal. In fact it was possible to detect the impurities in the space near the cathode with a direct-vision spectroscope, while in the positive column it was impossible to see any effect of impurities.

10. At pressures below 2 millimetres the positive column was no longer of uniform luminosity, and the force was found to depend on the current. At 1·4 millimetres pressure the force was 3·6 volts per centimetre for currents of about 8 milliamperes and 6 volts per centimetre for currents of about 2 milliamperes. At 1 millimetre pressure the force was 3 volts per centimetre for currents of about 8 milliamperes and 6·3 volts per centimetre for currents of about 2 milliamperes. These values were found both with the high frequency and direct discharge.

In some cases the direct discharge was unstable at pressures between 1 and 2 millimetres, and the current changed suddenly from 8 milliamperes to a low value of about one milliampere. This was accompanied by an increase of the force in the positive column. Within this range of pressures the positive column appeared to flicker faintly, which was no doubt due to changes in the current. These results are also shown in the curves (fig. 4) in which the ordinates give the force in volts per centimetre and the abscissæ the pressure p of the gas in millimetres. In this diagram the upper curve gives the values of X and of $2A/\pi$ obtained with currents of about 2 milliamperes, while the lower curve gives the forces obtained with currents of 8 milliamperes. The upper curve is a continuation of the curve obtained for all currents at pressures greater than 2·5 millimetres and shown in fig. 3, but the lower curve appears to have a discontinuity at about 2·5 millimetres pressure.

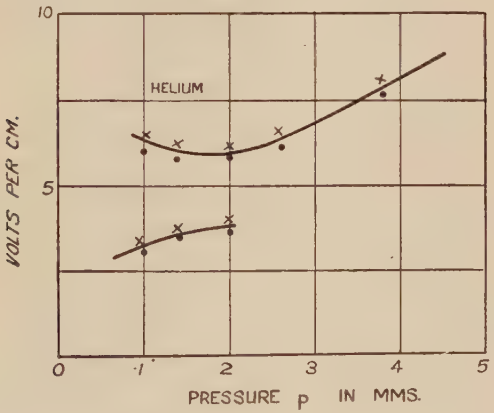
The discharges with the low currents were more stable than those with the higher currents.

Striations in the positive column of the direct current discharges began to be noticeable at pressures of about 2 millimetres, where slight differences of luminosity were observed in different sections of the tube. The brighter and darker sections recurred at equal intervals along the axis. At 1·4 millimetres the difference was more marked, and at 1 millimetre they were well defined. These striations took the form of bright ovals, about 3·5 centimetres long, with faint luminosity pervading the whole tube. The potential difference between the ends of

each striation was approximately 12 volts. These striations observed at low pressures were obtained with the purest specimens of helium. The high frequency glow did not appear to be striated at these pressures.

11. Striations were also observed in the high frequency discharge at high pressures when the gas was slightly impure. In the course of a few days any small amounts of impurity which may be emitted from the electrodes settle on the walls of the tube, but they can be driven off by heating. In one experiment pure helium at a pressure of 39 millimetres was admitted to the discharge-tube, which had not been heated for some days, and a uniform glow was obtained in the high frequency discharge. On heating

Fig. 4.



the tube gently, in order to drive off minute quantities of impurity, striations were observed. These striations consisted of a series of luminous disks, with their planes perpendicular to the axis of the tube. The potential difference across each striation was found to be 24 volts. When the discharge was stopped and the impurities absorbed by the charcoal in the side tube the striations were not reproduced.

Thus the disk type of striations only occurs in helium at high pressures in the presence of small amounts of impurity, while the oval striations occur at low pressures in the direct current discharge through the pure gas.

In conclusion, I wish to express my best thanks to Professor Townsend for his constant interest and advice.

XIV. *Radio Field-Strength Survey of the City of Calcutta and its Suburbs. By* Hrishikesh Rakshit, *M.Sc., Khaira Research Scholar in Physics, Calcutta University* *.

Introduction.

IT is well known that at short distances from a transmitter signals are received by the direct "ground-waves." The strength of the signals received locally does not show diurnal, nocturnal, or seasonal variation to any great extent, which is a marked feature of long-distance reception. On the other hand, the reception at short distances is considerably affected by the physical features of the locality, the presence of buildings and such other structures. Though it might appear at first sight that all places situated on the circumference of a circle with the transmitter at its centre should show equal field-strength, yet this is seldom met with in practice : two places situated at the same distance, but in different directions from a transmitter, may show considerable variation in the strength of the received signal. Thus every city with a broadcasting station has its own peculiar distribution of signal-strength. Local measurements of the strength of the signal in different directions supply a check on the laws of propagation of radio-waves, and also supply information about the suitable location of a broadcasting transmitter. Further, a knowledge of the values of field-strength in different directions measured at distances of a few wavelengths from the transmitter affords a means of calculating the effective height and radiation efficiency of the transmitting antenna.

Such measurements of local signal-strength have been made by Ralph Bown and G. D. Gillett †, Knox McIlwain and W. S. Thompson ‡. They have also drawn contour maps, showing lines of equal field-strength for New York, Washington (D.C.), and Philadelphia. The properties of Indian soil related to radio have never been studied before, and no contour map of signal-strength has been drawn for an Indian city. It was therefore considered

* Communicated by Prof. S. K. Mitra, D.Sc.
† Proc. I. R. E., Aug. 1924.
‡ Proc. I. R. E., Feb. 1928.

useful to carry out the measurements of signal-strength in and about the city of Calcutta, the method and results of which are given in the present paper. The observations were made with transmission from the local broadcasting station V.U.C. (λ=370·4 metres) with the co-operation of the transmitting staff, who maintained a constant power and wave-length during the hours of observation. The range of survey was an area swept by a radius of 12 miles round the transmitter.

Theoretical Discussion.

If a loop with sides very small compared with the wave-length of the incoming waves be tuned to resonance with the signals by means of a condenser, then the resonance voltage V developed across the tuning condenser is given by the equation

$$V = \frac{2\pi ANE}{R\lambda}\sqrt{R^2 + 4\pi^2 f^2 L^2} \text{ volts,}$$

where N = number of turns of wire in the loop,

A = area of each turn in sq. cm.,

E = electric field strength in volts per cm.,

R = total radio-frequency-loss-resistance of the loop-circuit in ohms,

λ = wave-length of the incoming waves in cm.,

f = frequency of the waves in cycles per sec., and

L = inductance of the loop in henrys.

For broadcast waves and for the loops generally used for reception $4\pi^2 f^2 L^2 >> R^2$, and we get

$$V = \frac{4\pi^2 ANE}{R\lambda} \cdot fL \text{ volts, approximately.}$$

If V and E be expressed in millivolts and millivolts/metre, then

$$E = V \times \frac{7{\cdot}6R}{f^2 LAN} \times 10^{10} \text{ m.v./metre.}$$

The values of f, A, and N are easily known; L, V, and R can also be measured without much difficulty; and from a knowledge of these E can be calculated.

Measurement of L.

The measurement of the inductance L of the loop was done according to the method adopted by Englund, Friis, and Bown *. The loop used was a square, 63 cm. on each side, and there were in all 12 turns, spaced at 8·5 mm. apart. The distributed capacity of the loop was found to be 16·65 $\mu\mu f$ and the inductance 180 μH.

Measurement of V.

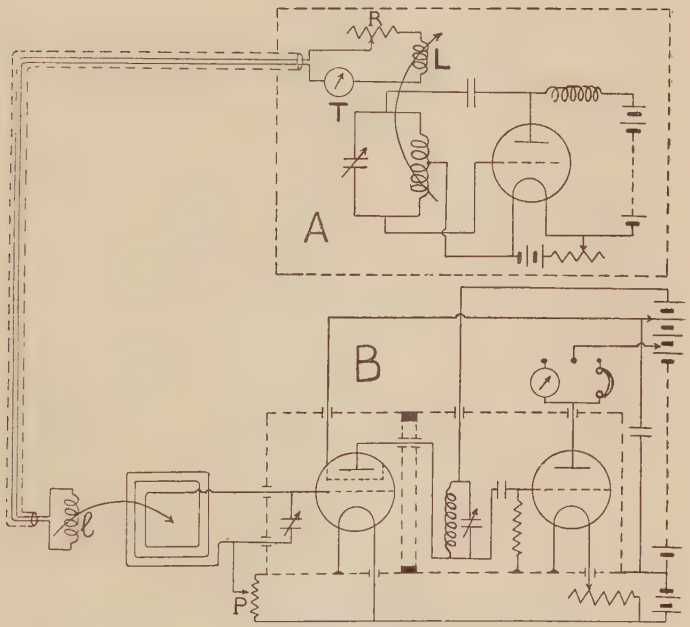
The principle of measuring the resonance voltage across the tuning condenser was a modification of that adopted by Knox McIlwain and W. S. Thompson †. Instead of a heterodyne receiver, a simple two-valve set was calibrated as a voltmeter in the laboratory. The receiver (B, fig. 1) consisted of one stage screen-grid tuned anode H.F. valve (voltage amplification factor=150) followed by a leaky-grid detector valve. One terminal of the loop-tuning condenser was connected to the grid of the H.F. valve and the other to the sliding contact P of a potentiometer connected across the filament battery. The grid, being thus at a potential higher than the negative end of the filament, effected a lowering of the resistance of the grid filament space. This resistance r, being in shunt with the tuning condenser C, was equivalent to a series resistance $1/4\pi^2f^2C^2r$. Though this caused a diminution of the sensitiveness of the receiver, yet the arrangement proved useful, as it checked the tendency to self-oscillation and also gave a control over the sensitiveness of the receiver.

The arrangement for calibration is shown in fig. 1. A carefully-screened oscillator A, placed at a certain distance from the receiver B, produced oscillations in the circuit containing a coil L, a variable non-inductive resistance R, a thermo-milliammeter T, and another small coil l, which was coupled to the frame aerial of the receiver, the coupling being kept constant throughout the calibration. The resonance voltage across the loop-tuning condenser for a known current in the coupling coil l, of the same frequency as that of the signals to be subsequently measured, was directly determined by a Moullin voltmeter. From this the resonance voltage across the loop-tuning condenser for any value of the current through l was easily calculated.

* Proc. I. R. E., April 1923.
† Proc. I. R. E., Feb. 1928.

The loop was then disconnected from the set and the steady current through the galvanometer in the plate-circuit of the detector noted. The oscillatory current through l being adjusted to a suitable value by means of R, the loop was again connected to the set, which was now carefully tuned by means of the loop-tuning and anode-tuning condensers, and the diminution of plate-current of the detector was recorded. Repeating this process for different values of current through l, the set was calibrated

Fig. 1.



Arrangement for calibrating the receiving set as voltmeter.

as a voltmeter, and a curve (fig. 2) was drawn. During the progress of the work the calibration was checked occasionally, and it was found to remain practically constant.

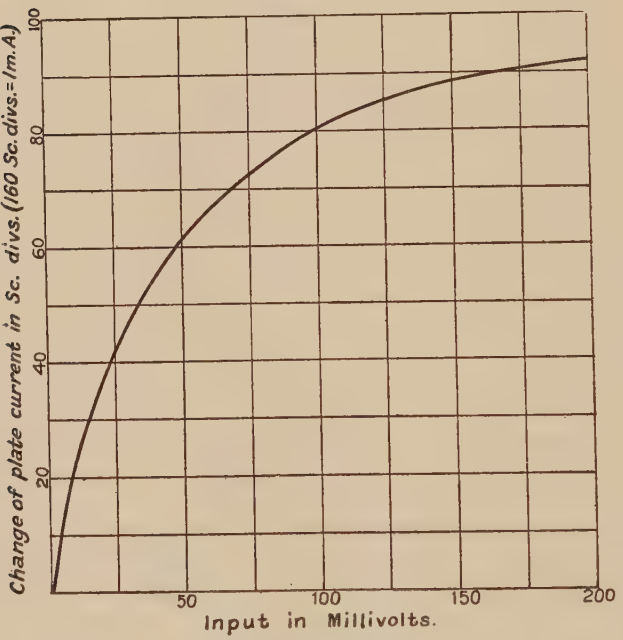
Measurement of R.

The resistance of the loop-circuit was measured by the usual reactance variation method *. During the hours of

* *Cf.* McIlwain and Thompson, Proc. I. R. E., Feb. 1928.

the day in which measurements were made in the field, there was no possibility of disturbance from any broadcasting station, and the only interference met on rare occasions, due to Morse signals, etc., was detected by the phones in the receiver. Owing to these facts, the measurements of loop-circuit resistance were made in the laboratory and only occasionally in the field, especially when inter-

Fig. 2.



Calibration curve.

ference was suspected. When the field-strength was sufficiently large a Moullin voltmeter was connected directly across the loop to measure the resonance voltage. In measurements with the Moullin voltmeter the shunt-resistance due to the grid-filament conductivity, as described previously, was absent. The resistance of the loop-circuit had thus obviously a different value, and it was therefore separately determined.

Description of a Typical Survey.

The car, with all the measuring equipment, was stopped at the place of observation, which was chosen beforehand with the help of a local map. In every case the car was stopped at a place as far as possible away from such obstacles as trees, large buildings, telegraph-posts and wires, overhead power-lines, etc., for these affected the field-strength considerably. The apparatus was set up, proper filament and high tension voltages applied, and the steady normal

Fig. 3.

H.F. plate voltage =80 ; screen grid voltage=60.
Det. plate voltage=40 ; fil. voltage (H.F.)=4.
Fil. voltage (det.)=2·4 ; potentiometer reading=4·5.
160 divns. of galvanometer=1 m.a.

1.	Location	Howrah Maidan
2.	Galvr. reading (loop off)	120
3.	Do. (loop on)	34
4.	Input in millivolts	130
5.	Galvr. reading to indicate half the resonance voltage	51
6.	(C_2-C_1) in $\mu\mu f$	28
7.	Resistance in ohms	35
8.	Direction	50° E.N.
9.	Field-strength (m.v/metre)	65

Typical Data Sheet.

reading of the galvanometer in the detector-plate circuit noted to ascertain that the set was in the condition similar to that obtained during calibration.

The set was properly tuned, the station listened-in by means of a pair of head-phones, and the loop rotated until maximum signal was heard. The presence of interference could then be easily detected, in which case further measurements were made only after the disturbance was over.

The phones were then replaced by the galvanometer, the readings of which were taken (*a*) with loop off, (*b*) with loop on (fig. 3, lines 2 and 3). From these the resonance

voltage (line 4) was determined from the calibration curve. The calibrated condenser across the loop was then rotated on both sides of the resonance value until half the resonance voltage was indicated, and from its values in the two positions the loop-circuit resistance was calculated (line 7).

The location and direction of loop being recorded, the batteries were disconnected, the galvanometer clamped, and the valves taken off their holders to prepare for transportation.

About 3 to 5 minutes were required to complete the observations at a single place when there was no interference, the presence of which required a longer period. Arriving at the laboratory, the field-strength was calculated and found to be 65 m.v. per metre (line 9) for the place.

Discussion.

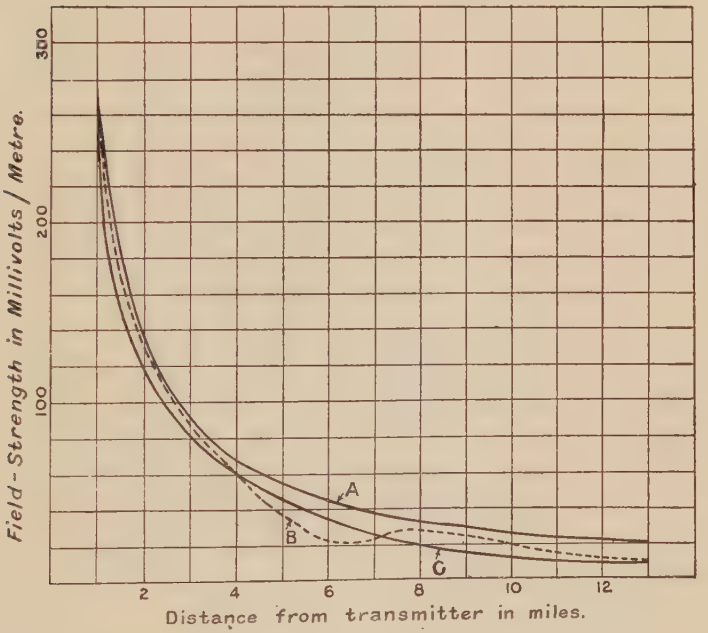
It is obviously very laborious to make close measurements in all directions, and for this reason it was decided to carry out such measurements in eight radial directions from the transmitter. All measurements were made, as far as possible, at places lying along these radial lines. After completion of the observations in one direction, a curve was drawn with field-strengths plotted against corresponding distances. Curve B (fig. 4) shows the result of measurement made in the south direction and curve C in the north-east direction. On drawing such a curve, some stray readings in the corresponding direction were usually taken to confirm the correctness of the curve, especially when it indicated abnormally high or low values of field-strength anywhere. Curve A shows the diminution of field-strength with distance if the waves simply spread out without being absorbed by the intervening media. From the average value of the field-strengths at about 2 miles from the transmitter in several directions, and from the knowledge of the aerial current of the transmitter, the effective height of the transmitting antenna was calculated, and curve A was drawn from this *. From a comparison of curves A and B it will be seen that in the south direction from about 3 miles from the transmitter the actual field-strength begins to diminish more rapidly than the

* *Cf.* Proc. I. R. E., April 1923, p. 117.

theoretical field-strength of curve A, reaching a minimum at a distance slightly greater than 6 miles. It then gradually regains its normal value, and finally falls off at normal rate. In the north-east direction (curve C) there is, however, no such anomalous minimum, but the field-strength diminishes continually, though it does not follow the theoretical curve.

With the help of curves like B and C for the eight directions, the contour map (fig. 5) was drawn. After

Fig. 4.



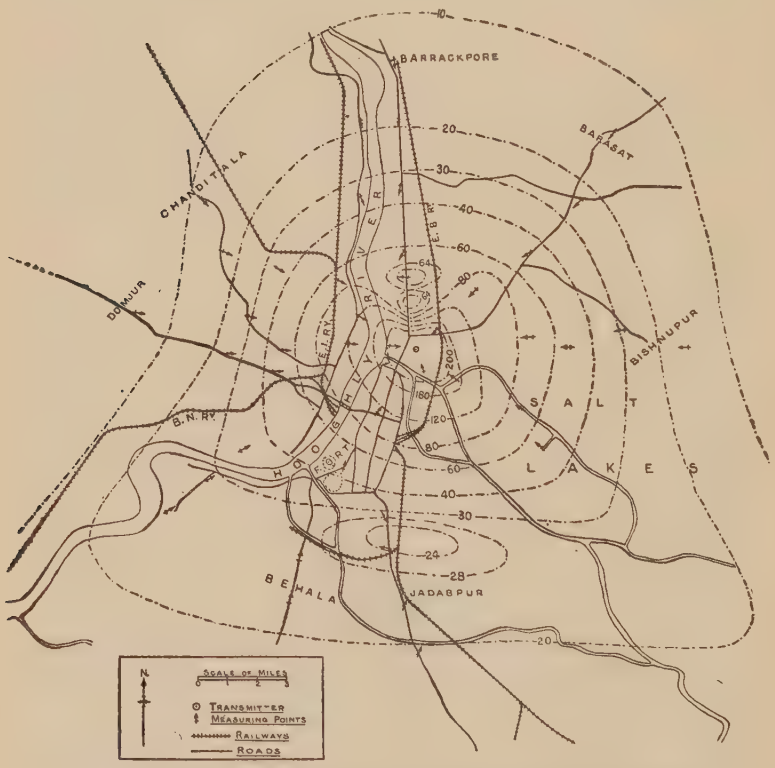
drawing the contour, a sufficient number of stray readings was taken in various directions to verify that the contour map so drawn represented the actual field-strengths at the places over which the contour lines passed.

The data given in this paper are the results of about two months (February and March 1930) of regular field-work, during which the car carrying the set travelled about 1000 miles, and about 200 readings were taken. The minimum field-strength measurable with the instrument used was

about 0·2 m.v. per metre, and a field-strength of 1 m.v. per metre could be measured with sufficient accuracy.

Fig. 5 shows the desired radio field-strength contour map of Calcutta and its suburbs. It will be seen that the contours roughly assume a circular shape with some exceptions. In the south-east direction there are vast lakes of salt water, and consequently the attenuation of

Fig. 5.



Radio field-strength contour map of Calcutta and its suburbs (field-strength in millivolts/metre).

the waves in this direction is very small, resulting in a bulging of the contour lines. In the south-west direction there is also a similar bulging, due to the River Hooghly. No appreciable bulging was, however, observed in the north direction due to the River Hooghly, although several readings immediately along the banks of the river were

taken to check this. At all the three positions of minima, *i.e.*, the centres of the closed loops, the apparent direction of arrival of the waves was nearly perpendicular to the direction of the transmitter, as indicated in fig. 5 by the directions of the arrows at the measuring points. From these anomalous directions of arrival of the waves it was concluded that at these places waves from different directions arrived and interfered, and thus caused the minima of field-strengths. On the south-south-west side of the transmitter, at a distance of 485 yards from it, there is a huge steel structure, height 110 feet, base 321 feet square, supporting a reservoir (Tallah Tank) of filtered water. Besides producing a very local shadow (not shown in fig. 5), it was found to have no effect on the general shape of the contours.

At a place 7 miles from the transmitter, and within the loop in the south direction, the field-strength was found to be as low as 0·9 m.v. per metre. This minimum effect was a persistent characteristic of the place, but so strictly local that it could not be shown on the map. The field assumed quite normal value at about 100 yards towards the west from the place of observation, but in the north and south directions the normal value was restored less rapidly. There was no large structure or any other obstacle near the location, and the apparent direction of arrival of the waves was nearly east-west, *i. e.*, at right angles to the direction of the transmitter. Measurements taken at this place during the night indicated that the field-strength approached very near the normal value at this distance, and the direction of arrival of the waves was from the transmitter. The position of minimum field, which was greater than that obtained during daytime, was, however, shifted by about 100 yards towards the transmitter, and the direction of arrival of the waves there was nearly east-west.

Conclusion.

The results of a quantitative radio field-strength survey of Calcutta and its suburbs have been given, and a contour map has been drawn showing lines of equal field-strength. It has been observed that the attenuation of the waves over land is much greater than that over water. The minima of field-strengths, which correspond to the loops in the contour map, have been found to be due to

interference between waves coming from different directions. It has been found that elevated structures cause local shadows, which are wiped out at larger distances due to energy coming from other portions of the wave-front. All these are in conformity with theoretical results, but, owing to the complexity of the media through which the radio-waves travel, the effect at any place cannot be rightly predicted *à priori* without the help of such contour maps.

The investigation was carried out under the direction of Prof. S. K. Mitra, D.Sc., to whom I tender my sincere thanks for constant help and guidance.

Wireless Laboratory,
University College of Science,
92 Upper Circular Road, Calcutta, India.
30th May, 1930.

XV. *A new Method of measuring the Capacity of Small Condensers.* *By* H. M. Barlow, *Ph.D.**

Introduction.

ONE of the best methods of measuring the capacity of a small condenser is to employ a commutator to charge and discharge it with a periodicity that can be accurately determined. If the average current in any one direction is measured by a galvanometer, and the applied potential is known, the capacity of the condenser may be readily calculated †.

When dealing with condensers under about 100 micro-micro-farads, the inherent capacity of the commutator and accessory apparatus becomes responsible for a considerable part of the total deflexion of the galvanometer, so that really accurate observations are not then possible.

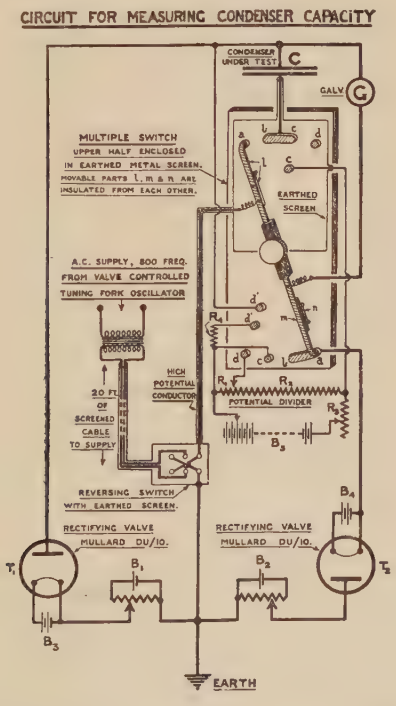
With the object of overcoming this difficulty an attempt was made to employ two thermionic valves in place of the commutator, the charging and discharging of the condenser being controlled by alternate positive and negative

* Communicated by the Author.

† J. A. Fleming and W. C. Clinton, Proc. Phys. Soc. xviii. p. 386 (1903), and Phil. Mag. v. p. 493 (1903).

potentials applied to the grids of the valves. This particular scheme proved unsatisfactory because commercial three-electrode valves have too large an impedance, except when the grids are very positive and taking large currents. Moreover, the high velocity with which the electrons in the valve are emitted from the hot filament cause a small current to flow round the plate circuit when

Fig. 1.


CIRCUIT FOR MEASURING CONDENSER CAPACITY
CONDENSER UNDER TEST
C
GALV
G
MULTIPLE SWITCH
UPPER HALF ENCLOSED IN EARTHED METAL SCREEN.
MOVABLE PARTS L, M & N ARE INSULATED FROM EACH OTHER.
EARTHED SCREEN
A.C. SUPPLY, 800 FREQ. FROM VALVE CONTROLLED TUNING FORK OSCILLATOR
HIGH POTENTIAL CONDUCTOR
20 FT. OF SCREENED CABLE TO SUPPLY
POTENTIAL DIVIDER
REVERSING SWITCH WITH EARTHED SCREEN.
RECTIFYING VALVE MULLARD DU/10.
RECTIFYING VALVE MULLARD DU/10.
EARTH

all the electrodes are at zero potential. This leads to an error, difficult to eliminate.

General Description of Apparatus employed.

The arrangement finally adopted consists in using the alternating potential itself to charge the condenser periodically in opposite directions through ordinary rectifying valves (fig. 1). The circuit is divided into two parts, so that the current for a positive charge is forced to

flow round one arm and for a negative charge round the other, an ordinary moving coil galvanometer G being used to indicate the magnitude of this current. The peak value of the applied alternating potential is determined by interrupting the valve circuit through T_2, thus maintaining the charge on the condenser C in one direction. A balance is then effected between the P.D. across the terminals of the condenser and the E.M.F. of the battery B_5, using the same galvanometer to establish this condition.

A system of resistances form a potential divider across the battery, and enable a simple calculation to be made for the capacity of the condenser in terms of these resistances and the frequency of the supply. A special multiple switch was constructed to make and break all the contacts in their proper sequence with four movements, *a*, *b*, *c*, and *d*, of a single handle. For simplicity in drawing this switch all the contacts are shown in the same plane, but in the actual article the upper and lower halves, which are electrically quite separate, were arranged vertically above one another to make the construction easier.

Batteries B_1 and B_2 were necessary to neutralize the electromotive forces due to the kinetic energy of the electrons emitted by the hot filaments of the Mullard DU/10 rectifying valves. In order suitably to adjust the compensation for this effect the multiple switch was set to position *a*, and one valve only, say T_1, with its attached neutralizing battery B_1, was put into operation. While the other valve T_2 was short-circuited, and the battery B_2 open-circuited, the potential injected from B_1 was increased until the deflexion of the galvanometer was just reduced to zero. A similar adjustment was made for the other valve T_2, so that neither taken separately acted as a source of E.M.F., but still had a low impedance to a current from plate to filament.

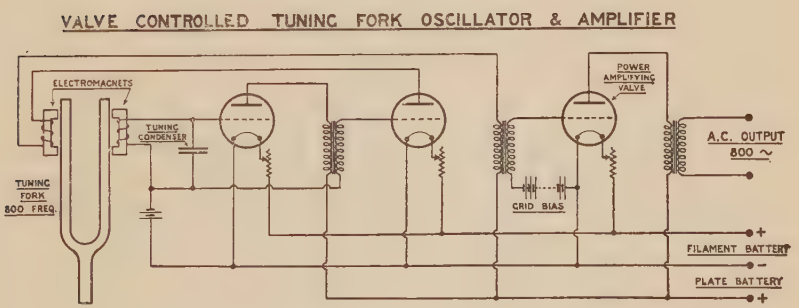
In some tests a new type of B.T.H. mercury-vapour valve with a hot cathode was used in place of the Mullard DU/10, and this was found to be more satisfactory under really exacting conditions, because its impedance was smaller, the neutralizing batteries B_1 and B_2 were not required, and "back-lash" current, from which most mercury-vapour valves tend to suffer, was completely absent.

A valve-controlled tuning-fork oscillator, having a frequency of 800, was employed to generate the alternating

current, and this was amplified before passing to the condenser testing circuit (fig. 2). To eliminate interference by stray electrostatic fields the oscillator and amplifier were separated from the rest of the apparatus by about 20 ft., one side of the A.C. supply was earthed, and all high potential conductors were very carefully screened. A reversing switch was inserted to enable a suitable correction to be made when dealing with an unsymmetrical voltage wave-form.

In some cases it is desirable also to employ a special intermediate circuit for flattening the crests of the wave and give the condenser more time to become fully charged (fig. 3). This is readily accomplished by using another

Fig. 2.

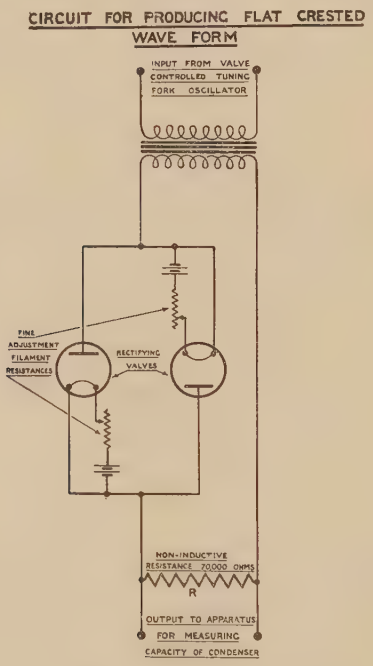


pair of rectifying valves, which quickly saturate as the applied potential rises *. With one valve operating on each half of the wave, the current passing through the system remains constant over a considerable part of the half period. Thus the wave-form of the P.D. across the non-inductive resistance R which forms the bulk of the load is more or less rectangular. The symmetry of the wave can be conveniently controlled by means of the filament rheostats. The valves used were Mullard Ora with grid and plate joined together. Their saturation current was found to be very suitable with applied potentials of about 200 volts.

* A device of this kind was employed to obtain a suitable time-scale for a cathode ray oscillograph by Rogowski and Glage, *Archiv. für Elektrotechnik*, ix. p. 115 (1920–21).

When an alternating E.M.F. is applied to the circuit shown in fig. 1 no current will flow into the condenser under test C if the multiple switch is in the position *a*. The high potential conductor connected to the movable arm *l* is then isolated, and the only current taken from the A.C. supply will be to satisfy its own capacity to earth. A corresponding return current flows in the earthed screen,

Fig. 3.


CIRCUIT FOR PRODUCING FLAT CRESTED WAVE FORM
INPUT FROM VALVE CONTROLLED TUNING FORK OSCILLATOR
FINE ADJUSTMENT FILAMENT RESISTANCES
RECTIFYING VALVES
NON-INDUCTIVE RESISTANCE 70000 OHMS
R
OUTPUT TO APPARATUS FOR MEASURING CAPACITY OF CONDENSER

but this does not affect the galvanometer. Turning the multiple switch into the position *b* joins the high potential conductor directly to the condenser, so that it is alternately charged in opposite directions. A series of unidirectional current impulses pass through each valve, and the average rate of flow through the tube T_2 is indicated by the galvanometer G. It should be observed that the whole of the high potential conductor to the actual point of contact with the condenser plate is carefully screened, so that the current returning *via* the galvanometer G is due entirely to

the charge on the condenser. No correction whatever is therefore required for the capacity of the measuring apparatus. Although not quite at earth potential, the galvanometer is never much above it, because the impedance of the valves is relatively negligible.

Having observed the galvanometer deflexion when the multiple switch is in position *b*, it is then moved into position *c*. The circuit through the valve T_2 is thus interrupted, and the current through the valve T_1 tends to maintain a negative charge on the low potential plate of the condenser. The battery B_5 is simultaneously connected at its negative end to the galvanometer and at its positive end through the variable resistance R_3 to the high potential plate of the condenser. If the voltage drop across the battery potential divider exceeds the P.D. maintained between the condenser plates, then the galvanometer will show no deflexion. By means of tapping points on the battery and the resistance R_3 the voltage across the potential divider is gradually reduced until it is equal to the crest value of the alternating E.M.F., a condition shown by the galvanometer just beginning to deflect. If the alternating E.M.F. has equal positive and negative crest values this balance will not be disturbed when the reversing switch in the A.C. supply is changed over. Making use of this fact, a symmetrical wave is readily obtained by adjusting the grid bias of the oscillator output valve. It is important that the battery B_5 and its attached resistances should be well insulated from earth, particularly at the positive end, to make leakage currents negligible.

The final movement of the multiple switch to the position *d* disconnects the A.C. supply and forms a local circuit through R_4 to the galvanometer and back to the tapping point between R_1 and R_2 on the potential divider (see also fig. 4). This tapping point is adjusted so that the galvanometer shows the same deflexion as it did when carrying current from the condenser in switch position *b*. The subsidiary contacts *d'* are only connected together in this final movement. Otherwise it was found that the capacity to earth of the battery B_5 and the various resistances might cause a slight error when the switch was in position *b*, especially with valves of relatively high impedance.

It is obviously not necessary for any of the resistances R_1, R_2, R_3, or R_4 to be non-inductive.

Theory and Method of Calculating Capacity of Condenser.

It is always preferable to use an alternating E.M.F. of symmetrical wave-form, or at least one having equal positive and negative crest values.

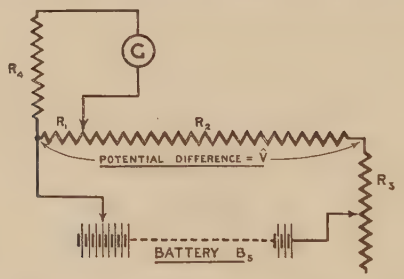
It is possible, however, to obtain perfectly accurate results with an alternating E.M.F. wave whose positive crest differs considerably in magnitude from the negative crest.

(a) *Symmetrical Alternating E.M.F. Wave-form.*

Consider first the case of a symmetrical wave-form whose positive and negative crest values are given by $\hat{V}$. With the multiple switch in position *b* the charge on the

Fig. 4.

GALVANOMETER CIRCUIT FOR MULTIPLE SWITCH IN POSITION "d"



condenser changes every half period from $+C\hat{V}$ to $-C\hat{V}$, so that, if the frequency is f, the average current through the galvanometer will be

$$2fC\hat{V}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (i.)$$

With the multiple switch in position *d* (see fig. 4) the galvanometer shows the same deflexion, and since the P.D. across R_1 and R_2 has already been adjusted to $\hat{V}$, it is easy to show that the galvanometer current is given by

$$\frac{R_1\hat{V}}{R_2(R_4+G)+R_1(R_4+R_2+G)}, \quad . \quad . \quad . \quad (ii.)$$

where G = galvanometer resistance.

Thus

$$2fC\hat{V} = \frac{R_1\hat{V}}{R_2(R_4+G)+R_1(R_4+R_2+G)}$$

or

$$C = \frac{R_1}{2\pi f[R_2(R_4+G)+R_1(R_4+R_2+G)]} \quad . \quad . \quad \text{(iii.)}$$

If the resistances are measured in ohms and the frequency in cycles per sec., then the capacity is given in farads. It is convenient to use relatively large values for the resistances R_4 and R_2, so that the denominator in the above expression remains practically constant, the term $R_1(R_4+R_2+G)$ being negligible, and the capacity becomes directly proportional to R_1.

To make an accurate calculation for C when R_1 is of the order of one per cent. or more of R_2 the complete expression must be used.

Thus, taking $f=800$,

$R_4=100{,}000$ ohms,

$R_2=10{,}000$ ohms,

and $G=730$ ohms,

it was found that a Standard Sullivan variable air condenser set to 60 micro-micro-farads gave balance when R_1 was 97·75 ohms. This agrees exactly with the theory. The value of R_3 was also made comparatively large, so that, using a resistance box for R_1, it could be varied within the required limits without appreciably disturbing the potential balance made when the multiple switch was in position *c*.

(b) *Unsymmetrical Alternating E.M.F. Wave-form.*

In this case let the positive crest value of the applied alternating E.M.F. be $\hat{V}'$, and the negative crest value $\hat{V}''$.

The current through the galvanometer from the condenser is given by

$$2\pi fC(\hat{V}'+\hat{V}'') \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(iv.)}$$

This reading is obviously unchanged by throwing over the reversing switch in the A.C. supply.

But there will be two different values of R_1 according to whether the balance is made against the positive or negative crest of the A.C. supply.

Thus, suppose R_1' corresponds with $\hat{V}'$,

and R_1'' corresponds with $\hat{V}''$,

both R_1' and R_1'' give the galvanometer deflexion obtained from periodically charging the condenser in opposite directions.

Hence

$$fC(\hat{V}' + \hat{V}'') = \frac{R_1'\hat{V}'}{R_2(R_4+G)+R_1'(R_4+R_2+G)} = \frac{R_1''\hat{V}''}{R_2(R_4+G)+R_1''(R_4+R_2+G)}; \quad . \quad (v.)$$

from which

$$C = \frac{K}{f[R_2(R_4+G)+2K(R_4+R_2+G)]}, \quad . \quad (vi.)$$

where

$$K = \frac{R_1'R_1''}{R_1'+R_1''} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (vii.)$$

For a symmetrical wave-form $R_1'=R_1''=R_1$, giving $K=\frac{R_1}{2}$, and expression (vi.) reduces to expression (iii.).

As an example of what may be done with an alternating E.M.F. having unequal positive and negative crest values, the following figures are given:—

$f=800$,
$R_4=100{,}000$ ohms,
$R_2=10{,}000$ ohms,
$G=730$ ohms.

For a capacity of 60 micro-micro-farads balance was obtained with $R_1'=53{\cdot}5$ ohms and $R_1''=557$ ohms. This gives $K=48{\cdot}8$.

Although the wave is completely unsymmetrical, it is easily seen that these figures lead to an accurate value of the capacity.

Applications and Limitations of the Method.

This method lends itself particularly well to the measurement of very small capacities. When using a reflecting galvanometer there is no difficulty in getting scale readings at a metre radius of as many centimetres as the capacity of the condenser under test. This is true at a frequency of 800, and there is no reason why higher frequencies, such as

those derived from quartz oscillators, should not be employed, particularly if the circuit for producing a flat-crested wave-form is incorporated. The limitations in this respect are set by the time constant of the charging or discharging circuit becoming comparable with the time during which the alternating E.M.F. may be considered to remain constant in any one direction. Using a slight flattening of the wave-form it was found possible to connect a non-inductive resistance of 100,000 ohms in series with either the valve T_1 or T_2 without introducing an error of more than 1 per cent. when measuring a capacity of 60 micro-micro-farads.

As already explained, the effect of the capacity of the measuring apparatus can be completely eliminated by suitable screening of the high potential conductor. The method is therefore capable of great accuracy, besides being very simple and easy to handle.

For condensers of 50 micro-micro-farads or more a unipivot galvanometer can be used in place of the reflecting type, and for a circuit with given components there is no difficulty in arranging matters so that the resistance R_1 is marked directly in micro-micro-farads. Like the Fleming and Clinton commutator arrangement it is not possible, for obvious reasons, to obtain an accurate measurement of the capacity of a really leaky condenser.

Moreover, one side of the condenser can never be quite at earth potential, because there must always be a small P.D. across the valves. This prevents the apparatus in the form described from being used to obtain the capacity to earth of an isolated conductor. The difficulty may be overcome by transfering the earth point to the low potential side of the condenser, so that the screen becomes insulated from earth. This arrangement is difficult to work. It is only practicable when the screen capacity to earth is reduced to a very small amount and exceedingly low impedance valves are employed.

In conclusion, the author desires to thank Prof. W. C. Clinton for giving him the use of the apparatus and the facilities of the Pender laboratories.

University College, London.
September 1930.

XVI. *On the Distribution of Electric Force and Rise of Temperature in the Glow Discharge.* *By* JOHN M. HOLM, *B.Sc., Strang Steel Scholar in the University of Glasgow* *.

(1) *Introduction.*

THE measurements described in the present paper were undertaken with the object of determining the simultaneous distributions of electric force and rise of temperature in a discharge-tube containing air or hydrogen at reduced pressure through which an electric current is passing. It is generally assumed that the distribution of temperature in such a discharge corresponds with the distribution of the rate of dissipation of energy per cubic centimetre, and therefore with the distribution of electric intensity ; but, as far as the present writer is aware, no accurate observations have ever been made with the object of testing the validity of the assumption. The matter is of some importance in connexion with existing theories of electrical ignition.

In a recent paper † Messrs. Finch and Cowen describe determinations which they had made of the rate of steam formation occurring in a steady glow discharge passing through electrolytic gas. Their conclusions are briefly as follows :—With small currents the combustion is confined to the immediate neighbourhood of the cathode, and the rate of such combustion, c_1, is

(*a*) independent of the pressure of the electrolytic gas,

(*b*) independent of the potential difference between the electrodes,

(*c*) independent of the gap-width,

(*d*) dependent on the nature of the cathode material, and

(*e*) is proportional to the current carried by the discharge.

When the current exceeds a certain value, i_2, combustion also takes place in the interelectrode zone, and the rate of this combustion, c_2, is :

(*f*) dependent on the pressure of the electrolytic gas,

* Communicated by Prof. E. Taylor Jones, D.Sc.

† Proc. Roy. Soc. A, cxi. p. 257 (1926).

(*g*) dependent on the gap width,

(*h*) independent of the material of the electrode, and

(*i*) is proportional to $(i-i_2)$, where i is the current carried by the discharge.

From these results the authors concluded that "the combustion which occurred was primarily determined by the ionization of the gaseous medium through which the current passed."

It has occurred to the present writer that the results of Messrs. Finch and Cowen might equally well be regarded as leading to the conclusion that the rate of combustion was proportional to the rate of dissipation of energy in the discharge. For, if c_1 is the rate of combustion near the cathode (cathodic combustion), then result (*e*) can be expressed as

$$c_1 = ai, \quad \ldots\ldots\ldots \quad (1)$$

where a is a coefficient which must satisfy conditions (*a*) to (*d*).

The normal cathode fall in potential, V_n, is known to conform to conditions (*a*) to (*d*), so that equation (1) can be written as

$$c_1 = \alpha V_n i, \quad \ldots\ldots\ldots \quad (2)$$

where α must fulfil conditions (*a*) to (*c*), but may or may not be independent of the cathode material. Since $V_n i$ is proportional to the total rate of dissipation of energy in the neighbourhood of the cathode, equation (1) is equally consistent with the view that the rate of combustion is proportional to the rate of dissipation of energy.

On this view the ratio of the rates of cathodic combustion for two different metals, say platinum and copper, should be

$$\frac{{}_{(\mathrm{Pt})}c_1}{{}_{(\mathrm{Cu})}c_1} = \frac{\alpha_{(\mathrm{Pt})}}{\alpha_{(\mathrm{Cu})}} \frac{{}_{(\mathrm{Pt})}V_n}{{}_{(\mathrm{Cu})}V_n},$$

where ${}_{(\mathrm{Pt})}c_1$ is the rate of combustion occurring near a platinum cathode.

Messrs. Finch and Cowen found that the value of this ratio was about 1·45, but their experiments did not include determinations of the normal cathode potential fall for these metals. However, the value 1·45 is probably too large for the ratio $\frac{{}_{(\mathrm{Pt})}V_n}{{}_{(\mathrm{Cu})}V_n}$, and the coefficient α must thus be assumed

to depend on the cathode material. The present writer would suggest that α depends on the cooling effect exerted by the cathode on the gas in the cathode zone, and diminishes when the thermal conductivity of the cathode material is increased. This would lead to a value of $\frac{(\mathrm{Pt})^{\alpha}}{(\mathrm{Cu})^{\alpha}}$ which is greater than unity.

With regard to the combustion in the interelectrode zone, let c_2 be the rate of combustion occurring in this region; then we have, in accordance with the conclusions of Messrs. Finch and Cowen,

$$c_2 = b(i - i_2), \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

where b is a coefficient which must satisfy conditions (f) to (h). The authors, however, offer no explanation of the variation of b with pressure and gap-width.

The present writer would suggest that the rate of combustion taking place in the interelectrode zone is proportional to the excess of the *rate of dissipation of energy* in that zone above the minimum rate of dissipation required for combustion, *i.e.*,

$$c_2 = \beta\{(\mathrm{V} - \mathrm{V}_n)i - (\mathrm{V}_2 - \mathrm{V}_n)i_2\}, \quad . \quad . \quad . \quad (4)$$

where V is the potential difference between the electrodes when the current carried by the discharge is i, and V_2 and i_2 are the values of these two quantities when combustion commences in the interelectrode zone.

From the graphs published by Messrs. Finch and Cowen (*loc. cit.*) it may be seen that V is practically independent of i for values of i equal to or greater than i_2. Thus equation (4) can be reduced to

$$c_2 = \beta(\mathrm{V} - \mathrm{V}_n)(i - i_2),$$

which is of the same form as equation (3), since $\beta(\mathrm{V} - \mathrm{V}_n)$ is practically constant in one series of experiments and fulfils conditions (f) to (h).

Table I. has been calculated from the results given by Messrs. Finch and Cowen for copper electrodes, and contains the values of β given by the ratio of the rate of combustion taking place in the interelectrode zone at a standard current of 3 milliamperes to the excess of the rate of dissipation of energy at that current above the minimum rate of dissipation required for combustion. The rate of combustion was evaluated in cubic centimetres of electrolytic gas at N.T.P. per minute, and the rate of dissipation of energy in milliwatts.

The value of ${}_{(Cu)}V_a$ was assumed to be 350 volts. p is the pressure of the electrolytic gas in millimetres of mercury, and d the gap-width in millimetres.

The average of the values of β obtained for the smallest gap-width is smaller than that for the other gaps. This is to be expected when the cooling effect of the anode on the interelectrode zone is considered, the effect being most pronounced with the smallest separation of the electrodes. The range of conditions represented in Table I. is very large; for example, the rate of combustion occurring in the interelectrode zone at the standard current of 3 milliamperes was ·035 c.c. at N.T.P. per minute with the 1·8 millimetre gap and a pressure of 30 mm. of mercury, while the rate at the same current was 4·25 c.c. at N.T.P. per minute with the 15 millimetre gap and a pressure of 75 mm. of mercury.

Table I.

p in mm.	d=1·8 mm.	d=3·8 mm.	d=7·5 mm.	d=15 mm.
30	·0269	·0307	·0341	·0377
45	·0285	·0320	·0320	·0370
60	·0252	·0285	·0305	·0330
75	·0265	·0280	·0260	·0365
90	·0267	·0306	·0266	--

When the above facts are taken into consideration, the approximate constancy of β is quite remarkable.

In a later paper* Messrs. Finch and Cowen describe experiments made to investigate the variation with pressure of the "minimum current" required to explode electrolytic gas. The discharge took place between water-cooled platinum electrodes.

The results obtained are represented by graphs showing the variation with pressure, for a fixed gap, of the "minimum current" required to explode the gas. The authors came to the conclusion that the graphs were hyperbolæ given by

$$p(i_3+1{\cdot}5)=c,$$

where p is the pressure of the electrolytic gas in centimetres of mercury, and i_3 is "minimum current" required to explode

* Proc. Roy. Soc. A, cxvi. p. 529 (1927).

the mixture at that pressure in milliamperes. The values of c were found to be 105, 68, and 55 for gap-width of 3·8, 7·5, and 15 millimetres respectively. The actual results, however, deviate considerably from these hyperbolæ at the lower pressures.

Messrs. Finch and Cowen concluded " that this proves that ignition is determined solely by the attainment in some portion of the gas traversed by the discharge of a certain definite concentration of suitable ions or electrically charged particles, because on kinetic ground the concentration of ions varies in a hyperbolic manner with the gas pressure. The curve cannot be accounted for on any purely thermal grounds because of the complete absence of lag, ignition invariably taking place immediately on attainment of the igniting current."

The present writer does not find these arguments convincing, and suggests that the results can be explained on the thermal theory advanced by Professor Taylor Jones *, and by Taylor Jones, Morgan, and Wheeler †, in connexion with ignition by electric sparks.

According to this theory no lag of any appreciable duration would be expected with water-cooled electrodes. There would be such a lag only if the running of the discharge for some time caused the electrodes to heat up and thus allow the temperature of the gas in the discharge to rise to its ignition value. This was prevented by the use of water-cooled electrodes. In their first paper (*loc. cit.*) Messrs. Finch and Cowen observed that, with the pointed wire electrodes then used, there was a lag of a few seconds between the attainment of the igniting current and ignition. It was found that this lag increased when the current was diminished. Both of these observations are in accordance with the thermal theory.

Again, according to this theory, for a given gap-width and pressure, the " minimum current " required to explode the mixture would increase when the electrodes were water-cooled. The numbers in Table II., obtained from the results published by Messrs. Finch and Cowen (*loc. cit.*), show that this is the case, and that the " minimum current " required to explode the gas, when water-cooled electrodes were used, was about *three times* that required with pointed wire electrodes of the same material—viz., platinum. The values of the " minimum current required for ignition " are expressed

* Phil. Mag. (December 1928).
† Phil. Mag. (February 1922).

in milliamperes. As before, p and d, the pressure of the electrolytic gas and gap-width, are given in millimetres of mercury and millimetres respectively.

The "minimum currents required for ignition" with the copper-wire electrodes are also included in Table II., and it may be seen that they are slightly greater than the corresponding values for platinum-wire electrodes, but much less than those for the water-cooled platinum electrodes. This is also to be expected on the thermal theory, since the cooling effect of the copper-wire electrodes on the surrounding gas will be greater than that of the platinum wire but much smaller than that of the water-cooled platinum electrodes.

The observed deviation from hyperbolæ of the graphs connecting the "minimum igniting current" with pressure

Table II.

	d=3·8 mm.			d=7·5 mm.			d=15 mm.		
p in mm.	Pt. Wire.	Cu. Wire.	Pt. Water-cooled.	Pt. Wire.	Cu. Wire.	Pt. Water-cooled.	Pt. Wire.	Cu. Wire.	Pt. Water-cooled.
45	—	—	—	—	—	—	2·7	>3·5	18
60	—	—	—	—	—	—	2	2·7	7·8
75	—	—	—	4	4·5	12	1·7	2·2	6
90	4·4	5·3	14	2·8	3	6·5	—	—	—

can also be explained on the thermal theory; for this theory, which contends that ignition begins on the attainment in some part of the discharge of a certain ignition temperature, would lead to the conclusion that, with water-cooled electrodes, the ignition would commence at a point practically midway between the electrodes. Messrs. Finch and Cowen also came to this conclusion (*loc. cit.* p. 536). The temperature attained at this point will depend primarily upon the rate of dissipation of energy, but will also be influenced by the distance of the point from the cold electrodes.

Assuming that the electric force in the positive column, Z, is uniform, then the rise in temperature, θ, at a point on the axis of the discharge midway between the electrodes will be given (see equation (17) below) by

$$\theta = AiZ, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

where A is a function which increases on increasing the gap-width and is independent of the pressure of the electrolytic gas.

Thus ignition will commence at this point when the product iZ reaches a certain value, given by

$$i_3 Z = \frac{\theta_1}{A},$$

where θ_1 is the difference between the ignition temperature of the gas and the temperature of the surrounding atmosphere.

At the pressures considered the electric force Z is nearly proportional to the pressure, *i. e.*,

$$Z = Bp,$$

where B is a function which diminishes slowly when the current is increased *. Hence

$$i_3 p = \frac{\theta_1}{A.B} = C \text{ (say)}, \quad . \quad . \quad . \quad . \quad . \quad (6)$$

where C is

(i.) independent of the pressure of the electrolytic gas,

(ii.) diminishes as the gap-width is increased, and

(iii.) increases slowly as the current is increased.

Equation (6), along with results (i.) and (iii.), show that the curve connecting the "minimum igniting current," i_3, with pressure, p, for a given gap-width will be approximately a hyperbola, but that it will deviate from a hyperbola in the required manner. Result (ii.) gives correctly the relative position of the curves for various gap-widths.

(2) *Experimental Arrangements.*

The purpose of the present experiments was to obtain direct evidence showing how the rise in average gas temperature along the path of a glow discharge varies with the rate of dissipation of energy. Assuming the current density across the section of the discharge constant, then the rate of dissipation of energy per unit volume is proportional to iZ where i is the total current flowing in the discharge and Z the longitudinal electric force. Hence determinations were

* *Handbuch der Physik*, Bd. xiv. S. 271.

made of the distributions of rise in temperature and longitudinal electric force along the axis of the discharge over as large a pressure range as possible. Both of these distributions have already been studied *, but, so far the present writer is aware, always in separate discharges.

The discharge-tube used was of a similar type to that employed by H. A. Wilson (*loc. cit.*). Electrodes of "commercial" aluminium, turned to be a sliding-fit inside the tube, were kept at a fixed distance apart, 14 cm., by a single thin glass rod, A (see fig. 1). Attached to the back of one of the electrodes was an iron disk, so that the electrode system could be drawn along the tube, which was horizontal, by an external magnet. Spirals of thin brass wire, B, were used to connect the electrodes with the ends of the tube. One end of the tube was closed by a brass cap, secured by a cement composed of beeswax and resin, so that the electrode system could be removed and the inner surface of the tube cleaned when coated with spluttered material. It was found that this operation had to be carried out much more frequently when the discharges took place in air than when they passed through hydrogen. At the middle of the discharge-tube, which was about 60 cm. long, two short, diametrically opposite side-tubes were attached, so that various probes could be inserted into the discharge. The tube was exhausted by an oil-pump which could be cut off from the discharge-tube by a large glass stopcock. To one limb of a T-tube, joined to the discharge-tube near one end, a manometer was attached, while the other limb gave access, through a stopcock, to a system for drying the gas used, and, in the case of hydrogen, to an electrolysis apparatus. The air used was passed through concentrated sulphuric acid and the hydrogen was prepared by electrolysis of a solution of barium hydroxide. The latter gas was dried by calcium chloride and phosphorus pentoxide.

An Evershed Wireless Generator, in series with one-hundred and twenty small accumulator cells, furnished the potential necessary for the maintenance of the discharge. The current from the generator, which had a maximum output of 30 milliamperes at 1000 volts, was smoothed by the usual condenser, choke, condenser circuit, the hum in a pair of sensitive headphones, when connected in series with one megohm across the output, being practically inaudible. Control over the current passing through the discharge was

* R. W. Wood, *Ann. der Phys.* Bd. lix. S. 238 (1896); H. A. Wilson, Phil. Mag. xlix. p. 505 (1900).

effected by adjusting the speed of the generator motor. A standard megohm in series with a milliamperemeter was utilized to determine the potential difference between the electrodes, while a Gaiffe milliamperemeter was used to measure the discharge current. The whole of the electrical circuit was insulated from the table by paraffin-wax blocks or thick rubber tubing.

(3) *The Electric Force.*

The method used was similar to that employed by H. A. Wilson (*loc. cit.*), and consisted in measuring electrostatically the difference in potential acquired by two insulated parallel probes placed in the discharge at a known small distance, δz, apart. The probes were so placed that they lay on a plane passing through the axis of the tube If $\delta V'$ is the difference in potential between the probes, then the electric force is $\frac{\delta V'}{\delta z}$.

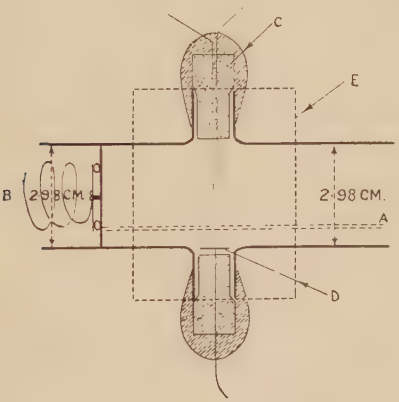
The difference in potential, $\delta V'$, was found by connecting the probes to a well-insulated mica condenser, allowing sufficient time for the condenser to charge up, and then discharging it by means of a double throw, double pole switch through a ballistic galvanometer. An interval of one and a quarter minutes was found to be sufficient for the condenser used, the latter having a capacity of approximately one-tenth of a microfarad. The sensitivity of the ballistic galvanometer was about 500 mm. per microcoulomb at 1·3 metres.

The method was considered to be superior to the usual quadrant electrometer method, since it eliminated the long leads from the probes to the electrometer, the condenser being mounted quite close to the probes. The electrometer method, in which both quadrants are "free," is open to another objection, for in the circumstances of the present experiments it cannot be assumed that the potentials of the quadrants are small in comparison with that of the needle, and, consequently, the deflexion of the needle cannot be taken as proportional to the difference of potential of the probes.

The probes finally used consisted of two pieces of platinum wire ·04 mm. in diameter, enclosed in glass capillary tubing ·15 mm. in external diameter, to within 1·5 mm. of their ends, and mounted in an ebonite holder, C, so that they were parallel to each other and about 1·3 mm. apart. Paraffin wax was used to fix the probes in the holder, which was cemented into one of the side-tubes (see fig. 1).

To test if the method was giving the electric force correctly, a small brass disk, D, was fitted into the other side-tube. This being used as an anode, the potential difference between it and the cathode was determined when the cathode was placed so that the disk D was just outside the negative glow, the current being adjusted till it had the same value as that which had been carried by the main discharge. This potential, V_1, was taken to be the sum of the cathode and anode falls in potential for the main discharge, the auxiliary anode D being covered with an anode glow. From the distribution of electric force which had previously been determined over the region extending

Fig. 1.



from the tail of the negative glow to near the anode the fall in potential between these two points was calculated. Adding to this the corresponding value of V_1, the result obtained was compared with V_2, the total fall in potential between the electrodes as measured by the external voltmeter—*i. e.*, the accuracy of

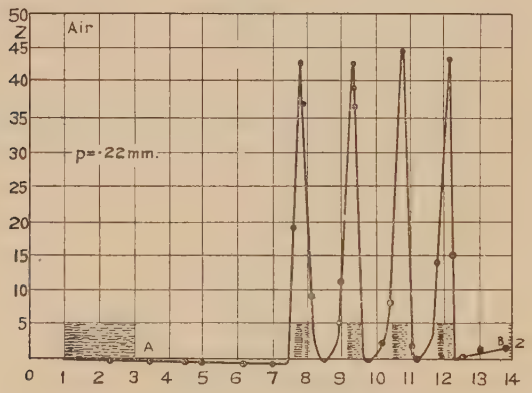
$$V_2 = V_1 + \int_A^B Z\,dz \quad . \quad . \quad . \quad . \quad . \quad (7)$$

was tested, A and B being points in the tail of the negative glow and head of anode glow respectively. The above is a necessary but not sufficient test as to the accuracy of the values of the electric force.

With discharges in air the validity of equation (7) was examined at pressures ranging from ·05 to 4 mm. of

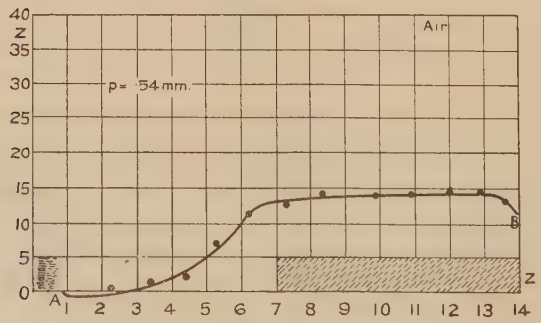
mercury, and with hydrogen at pressures from 2 to 8 mm. The various types of distribution of electric force obtained are shown in figs. 2 to 5. The types represented in figs. 3 and 4 were common to both air and hydrogen.

Fig. 2.



$i = \cdot 35$ ma. $V_2 = 648$ volts. $V_1 = 570$ volts. $\int_A^B Z\,dz = 67$ volts.

Fig. 3

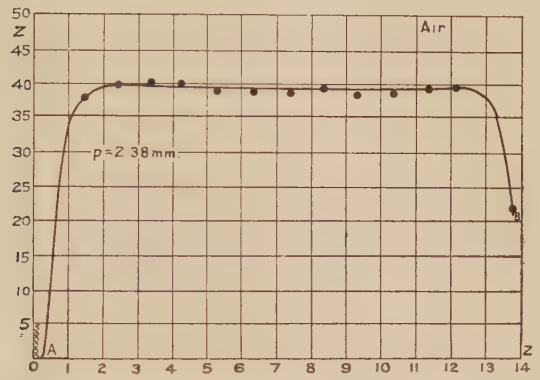


$i = 1{\cdot}25$ ma. $V_2 = 585$ volts. $V_1 = 435$ volts. $\int_A^B Z\,dz = 119$ volts.

In figs. 2 to 5 the abscissæ represent distances from the cathode in centimetres and the ordinates the values of the electric force, Z, in volts per centimetre. The distribution of light in the discharge is roughly indicated by the shading along the $0z$ axis.

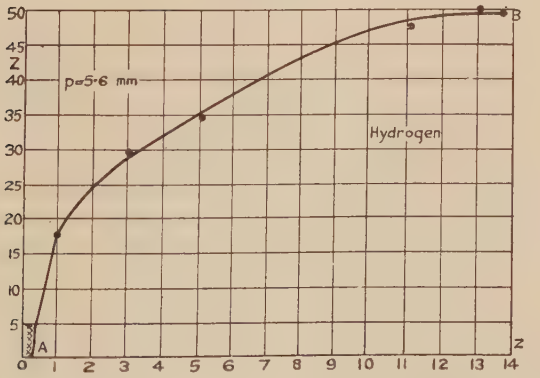
Striated Discharge (fig. 2).—In the case of the striated discharge the above test was not very exacting, on account of the small value of $\int_A^B Z\,dz$ and the large value of V_1.

Fig. 4.



$i = 2{\cdot}45$ ma. $V_2 = 840$ volts. $V_1 = 335$ volts. $\int_A^B Z\,dz = 504$ volts.

Fig. 5.



$i = 4$ ma. $V_2 = 822$ volts. $V_1 = 305$ volts. $\int_A^B Z\,dz = 513$ volts.

Since the discharge was strongly abnormal the value of V_1 was considerably altered by a small change in discharge current. For example, in the case of three striations the

value of V_1 was 572 volts with a current of ·30 ma. and 584 volts with a current of ·35 ma., while the value of $\int_A^B Z\,dz$ was 13 volts and of V_2 582 volts at ·30 ma.

However, the results obtained yield a quite possible value for $\int_A^B Z\,dz$, and are very similar to those obtained by Sir J. J. Thomson *, who used the deflexion of a beam of electrons to determine the electric field in the neighbourhood of hydrogen striæ. From the results published by W. H. McCurdy †, the value of $\frac{\partial \nu_0}{\partial z}$ (see equation (28) below) is seen to be zero at the head and also at the tail of striæ in mercury vapour. Assuming that this is also true for striations in air, then the maximum value, Z_m, of the electric force observed at the head of each stria will be correct. The value of $\frac{Z_m}{p}$ was found to be approximately constant for discharges containing one, two, three, and four striæ, and to have a value of approximately 167 volts per cm. per mm. of mercury, p being the pressure of the air in millimetres of mercury.

Bright Uniform Positive Column (fig. 3).—With this type of distribution the values of $V_1 + \int_A^B Z\,dz$ were from 5 to 10 per cent. lower than the values of V_2. This probably indicated that the values obtained for the electric force in the Faraday dark space were too small.

No Visible Positive Column (fig. 4).—In the type represented in fig. 4 the positive column, if visible, was very faint, and extended to within one or two centimetres from the cathode. The value of $V_1 + \int_A^B Z\,dz$ was found to agree with that of V_2 to within 1 per cent. The ratio of the constant value of the electric force $\bar{Z}$ to the pressure was found to diminish slowly with increasing pressure. In air the value of the ratio $\frac{\bar{Z}}{p}$ was 17·2 volts per cm. per mm. of

* Phil. Mag. p. 449 (1909).
† Phil. Mag. xlviii. p. 898 (1924).

mercury at a pressure of 1·7 mm., and 15·7 at 3·7 mm. With hydrogen discharges the value of $\frac{\overline{Z}}{p}$ was 10·9 at 2·3 mm., and 8·3 at 4 mm.

Fig. 5.—The distribution shown in fig. 5 was only obtained in the case of hydrogen, and was associated with the contraction of the usual anode glow into an intense ball or spot of light at the centre of the anode. The agreement between $V_1 + \int_A^B Z\,dz$ and V_2 was again within 1 per cent., the range of pressures examined being from 4 to 8 mm. of mercury.

Since it was difficult to keep the conditions of the discharge—i. e., i, V_2, and p—constant to within 1 per cent. over a set of readings, which usually occupied about half an hour, the agreement in the cases of figs. 4 and 5 is as good as could be expected.

At this point it might be convenient to refer to the appearance of a peculiar striation found in air discharges at pressures between 1 and 2 mm. of mercury. This took the form of a red stria of the same sectional area as the negative glow, and appeared in the Faraday dark space, close to the negative glow, being about 1·5 cm. from the cathode. It could only be obtained when the current carried by the discharge was reduced as far as possible, usually to about 1·5 milliamperes, and frequently appeared just before the discharge went out. The striation could be obtained either with or without a visible uniform positive column, and was separated from the head of the positive column, when the latter appeared, by a Faraday dark space about 6 centimetres in length. In one case two, and in another three such striæ were observed separated from each other by about 5 millimetres. The thickness of each striation was about 2 or 3 millimetres. Several determinations of the distribution of electric force in and near the isolated stria were made, and corresponded to the type shown in fig. 4. From these the difference in potential between the edge of the negative glow and the head of the stria was found to be about 25 volts. Thus it appears improbable that the striation is due to the first resonance collision of electrons starting with zero velocity at the edge of the negative glow, as the wave-length of the light emitted therefrom would be in the ultra-violet, and the potential difference between the edge of the negative glow and the striation would only be about

8 volts. Several photographs were taken of this peculiar red striation, and on some of these another striation was visible about halfway between it and the edge of the negative glow. Since the light emitted by this stria must be in the ultra-violet, and since the potential difference between the edge of the negative glow and its head would be very roughly about 8 volts, it is possible that it is due to the above cause.

The above phenomenon seems similar to that described by Prof. J. Zeleny *, who observed that in hydrogen discharges at a pressure of about 1·5 mm. of mercury by reducing the pressure slightly one (or more) striations broke off from the positive column and moved slowly across the Faraday dark space to near the negative glow. However, in the present case the striation was formed just outside the negative glow, and did not appear to have any connexion with the positive column.

(4) *The Rise in Temperature.*

A small platinum wire thermometer was constructed from Wollaston wire, the diameter of the platinum wire being stated by the makers to be ·005 mm., while that of the silver sheath was ·05 mm. The thermometer consisted of a hairpin-shaped loop of the wire, mounted in an ebonite holder similar to that used for the probes, and having the silver dissolved away just at the bend, leaving about 3 mm. of the platinum wire bare. This short piece of wire had a resistance of 18·673 ohms at 20° C. The value of the mean temperature coefficient of resistance between 20° and 100° C. was ·00284 per degree C., indicating that the platinum was not very pure. A Callender and Griffiths' Bridge was used to measure the resistance of the thermometer wire. With a constant current of ·7 milliampere flowing through the thermometer the sensitivity was such that a shift of 1 mm. in the point of balance on the bridge wire could be detected, corresponding to a change of temperature of ·0077° C. In order to shield them from draughts, the compensating leads, along with the actual leads to the thermometer wire, were enclosed in rubber tubing. The ebonite holder, to which these leads were attached, was inserted in the side-tube opposite the one containing the probes and secured with cement, so that the thermometer wire was about 1 millimetre from the axis of the discharge-tube. The

* 'Nature,' cxxv. p. 562 (1930).

thermometer lay in the same cross-section as one of the probes, and probably indicated the average gas temperature at about 2 mm. from the axis.

In the experiments made on the rise in temperature, a water-jacket, E, was fitted round the discharge-tube, as shown in fig. 1, to a distance of 2·3 cm. on either side of the section containing the probe and the thermometer wire. The jacket was filled with 100 c.c. of distilled water. Immediately before each reading of the temperature in the discharge was taken the temperature of the water was noted, and this was taken as t_0 in the equations of section (5). The time-rate of increase of t_0 was in good agreement with that calculated from equation (9). For example, in the discharge represented in fig. 7, where iZ was 154 milliwatts per centimetre, giving a calculated rate of rise of t_0 of ·112° C. per minute, the rate observed was ·11° C. per minute.

Figs. 6–8 represent the results obtained for the distribution of the rise in temperature along the path of the discharge as measured by the platinum-wire thermometer, and the corresponding distribution of the rate of dissipation of energy. As before, the abscissæ are distances from the cathode in centimetres. The ordinates of the iZ curves give the rates of dissipation of energy in the discharge expressed in milliwatts per centimetre, and those of the $(t'-t_0)$ curves the rise in temperature of the platinum thermometer wire in degrees centigrade.

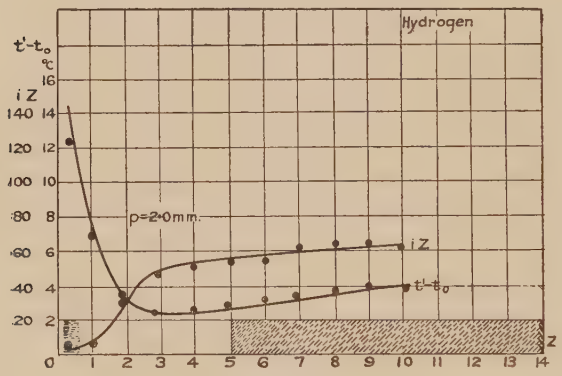
(5) *Discussion of Results.*

In all cases the graphs of $(t'-t_0)$ were practically parallel to those of iZ, except near the cathode, where they crossed and diverged from each other. This was due to the conduction of heat parallel to the axis of the tube, Oz, from the cathode zone into the Faraday dark space. The values of $\frac{i\mathrm{Z}}{(t'-t_0)}$ for figs. 6, 7, and 8, taken at $z=8$ cm., are 17·1, 13·0, and 15·0 milliwatts per centimetre per degree centigrade. From these results it is evident that the rise in temperature, as given by the resistance of the platinum thermometer, is not exactly proportional to the rate of dissipation of energy.

Before proceeding further it will be necessary to consider the theory of the rise in temperature in the discharge as calculated from the rate of dissipation of energy. This

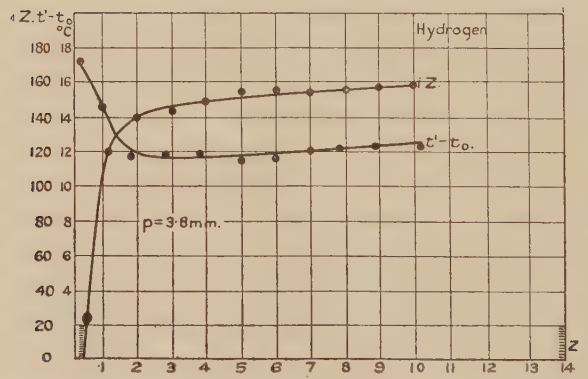
problem was first solved by Prof. E. Warburg *, but as it is of importance in connexion with the present work the required result will be derived in a slightly different manner.

Fig. 6.



$i = 3$ ma. $V_2 = 570$ volts.

Fig. 7.



$i = 4{\cdot}25$ ma. $V_2 = 800$ volts.

When expressed in cylindrical coordinates, r, θ, z, the general equation for the conduction of heat for a system in equilibrium is

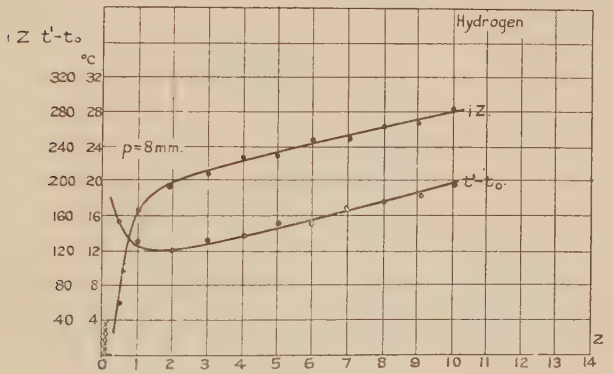
$$\frac{1}{r}\frac{\partial}{\partial r}\left(kr\frac{\partial t}{\partial r}\right)+\frac{1}{r^2}\frac{\partial}{\partial \theta}\left(k\frac{\partial t}{\partial \theta}\right)+\frac{\partial}{\partial z}\left(k\frac{\partial t}{\partial z}\right)+q=0, \quad (8)$$

* *Ann. der Phys.* Bd. liv. S. 265 (1895).

where t and k are the temperature and thermal conductivity of the system at the point (r, θ, z), and q is the quantity of heat produced at (r, θ, z) per second per unit volume. In equation (8) k is not assumed to be independent of r, θ, and z—*i. e.*, it may be a function of t.

Let us consider the application of equation (8) to the case of a glow discharge taking place in a cylindrical discharge-tube of inner radius R and outer radius $R + \delta$. Let the electrodes be circular metal disks of radius R, placed with their planes perpendicular to the axis of the tube, and kept at a fixed distance apart which is large compared with the radius of the tube. Let the outer surface of the discharge-tube be maintained at a constant temperature, t_0, by means of a

Fig. 8.



$i = 6{\cdot}1$ ma. $V_2 = 900$ volts.

water-jacket through which a steady stream of water is circulating. Assume that the discharge is of the uniform positive column type, in which the longitudinal electric force Z is constant over a considerable part of the discharge. Figs. 3 and 4 indicate the distribution of Z in suitable discharges.

Let the axes of coordinates be so chosen that Oz coincides with the axis of the tube and O, the origin, with the inner surface of the cathode. Let i be the total current flowing in the discharge in milliamperes. Assuming that Z and the current density parallel to Oz are independent of r and θ, then we have

$$q = \frac{10^4}{\pi R^2 J} . iZ \text{ cal. per c.c. per sec., . . . (9)}$$

where J is the mechanical equivalent of heat expressed in C.G.S. units, R the radius of the tube in centimetres, and Z the electric force in volts per centimetre.

If we consider a section of the discharge such as that near the middle of the visible positive column in fig. 3, or about halfway between the negative glow and the anode in fig. 4, then q is practically independent of r, θ, and z, and the flow of heat is therefore purely radial—*i. e.*, $\frac{\partial t}{\partial z}$ and $\frac{\partial t}{\partial \theta}$ are zero in equation (8).

Let k be the thermal conductivity of the gas; then for $0 \leq r \leq \mathrm{R}$ we have

$$\frac{1}{r}\frac{\partial}{\partial r}\left(kr\frac{\partial t}{\partial r}\right)+q=0. \quad . \quad . \quad . \quad . \quad (10)$$

Integrating equation (10), and using the condition that $-2\pi rk\frac{\partial t}{\partial r}=\pi \mathrm{R}^2 q$ when $r=\mathrm{R}$, we obtain

$$k\frac{\partial t}{\partial r}+\tfrac{1}{2}qr=0. \quad . \quad . \quad . \quad . \quad . \quad (11)$$

Now k is a linear function of the temperature*, and we may assume

$$k=k_0(1+\alpha t). \quad . \quad . \quad . \quad . \quad . \quad (12)$$

Substituting from equation (12) in equation (11), integrating and solving for $(t-t_{\mathrm{R}})$, we obtain

$$t-t_{\mathrm{R}}=\frac{-1+\sqrt{1+\frac{\alpha q}{2k_0}(\mathrm{R}^2-r^2)}}{\alpha}, \quad . \quad . \quad (13)$$

where t_{R} is the temperature of the inner surface of the discharge-tube.

In order to find t_{R} let k' be the thermal conductivity of the material of the discharge-tube, and let it be assumed constant. For $\mathrm{R} \leq r \leq (\mathrm{R}+\delta)$ q is zero, and equation (8) gives

$$\frac{1}{r}\frac{\partial}{\partial r}\left(k'r\frac{\partial t}{\partial r}\right)=0.$$

Integrating twice with respect to r, and using the conditions that $-2\pi rk'\frac{\partial t}{\partial r}=\pi \mathrm{R}^2 q$ when $r=\mathrm{R}$, and that

* A. Winkelmann, *Ann. der Phys.* Bd. xliv. S. 456 (1891).

$t = t_0 =$ temperature of the water in the water-jacket when $r = \mathrm{R} + \delta$, we have

$$t_{\mathrm{R}} = t_0 + \frac{\mathrm{R}^2 q}{2k'} \log_e \left(\frac{\mathrm{R}+\delta}{\mathrm{R}} \right). \quad . \quad . \quad . \quad . \quad . \quad (14)$$

Thus, for $0 \leq r \leq \mathrm{R}$,

$$t = t_0 + \frac{\mathrm{R}^2 q}{2k'} \log_e \left\{ \frac{\mathrm{R}+\delta}{\mathrm{R}} \right\} + \frac{-1 + \sqrt{1 + \frac{\alpha q}{2k_0}(\mathrm{R}^2 + r^2)}}{\alpha}.$$

$$. \quad . \quad . \quad (15)$$

The term $\frac{\alpha q}{2k_0}(\mathrm{R}^2 - r^2)$ can be shown to be small compared with unity in the case of a hydrogen discharge, so that, as a first approximation, we obtain

$$t - t_0 = \frac{1}{2} \left\{ \frac{1}{k'} \log_e \left(1 + \frac{\delta}{\mathrm{R}} \right) + \frac{\left(1 - \frac{r^2}{\mathrm{R}^2} \right)}{2k_0} \right\} \mathrm{R}^2 q. \quad (16)$$

Substituting from equation (9) for q, and assuming $\frac{\delta}{\mathrm{R}}$ small compared with unity, equation (16) reduces to

$$t - t_0 = \frac{10^4}{2\pi \mathrm{J}} \left\{ \frac{\delta}{k' \mathrm{R}} + \frac{1}{2k_0} \left(1 - \frac{r^2}{\mathrm{R}^2} \right) \right\} i\mathrm{Z}. \quad . \quad (17)$$

Thus the rise of temperature in the type of discharge considered is, to a first approximation, proportional to the rate of dissipation of energy.

In the experiments described in section (2) the values of R and δ were 1·49 and ·13 cm. respectively. The discharge-tube was made of soft glass tubing, so that k' could be assumed to be ·0018 per degree centigrade. Assuming that the value of k_0 for hydrogen is ·000344 *, and that $\mathrm{J} = 4{\cdot}185 \cdot 10^7$ ergs per calorie, then equation (17) gives

$$t - t_0 = {\cdot}0000380 \left\{ 48{\cdot}5 + 1453 \left(1 - \frac{r^2}{\mathrm{R}^2} \right) \right\} i\mathrm{Z}. \quad (18)$$

Thus the rise in temperature at the thermometer wire, *i. e.*, at $r = {\cdot}2$ cm., will be

$$t_{\mathrm{Pt}} - t_0 = {\cdot}0562\, i\mathrm{Z}. \quad . \quad . \quad . \quad . \quad . \quad (19)$$

Hence

$$\frac{i\mathrm{Z}}{t_{\mathrm{Pt}} - t_0} = 17{\cdot}8. \quad . \quad . \quad . \quad . \quad . \quad (20)$$

* Meyer, 'Kinetic Theory of Gases,' p. 295.

As stated above, the values of this ratio as obtained from figs. 6, 7, and 8 were 17·1, 13·0, and 15·0 respectively. Thus the temperature, as determined by the resistance of the platinum wire, is too high in all cases. The actual differences between the observed rise in temperature and that calculated from equation (19) are ·12° C., 3·5° C., and 2·7° C. respectively.

This discrepancy is too large to be explained either by experimental error or by incorrect values chosen for the constants in equation (17). Allowance for the conduction taking place along the leads of the platinum-wire thermometer, or inclusion of the third term in the expansion of the square root in equation (15), would both reduce the calculated value of the rise in temperature. The present writer would suggest that the difference between the rise in temperature indicated by the platinum-wire thermometer and the calculated rise is due to the bombardment of the wire by positive ions accelerated towards it and to the subsequent neutralization of the positive ions by electrons on the surface of the wire.

To calculate the magnitude of this effect let n_2 positive ions reach the thermometer wire per unit length per second, and let the latter be at a potential which is v volts below the actual potential of the surrounding discharge. Then the energy liberated on the surface of the wire per second, due to the positive ion bombardment, is $n_2 e v$, where e is the charge on an electron.

Now the positive ions arriving at the surface of the wire must be neutralized there by the equivalent electron current flowing to the surface of the wire along the thermometer leads. This will also liberate energy on the surface of the wire. Let V_i be the ionization potential of the positive ions, then the amount of energy liberated due to this cause will be $n_2 e V_i$ per unit length of the wire per second.

Let q' be the total quantity of heat liberated on the surface of the thermometer wire, *i. e.*,

$$q' = n_2 e v + n_2 e V_i = (v + V_i) n_2 e.$$

Let I_+ be the positive ion current per unit length of the wire in milliamperes ; then

$$q' = \frac{10^4}{J} (v + V_i) I_+ \text{ cal. per cm. per sec.,} \quad . \quad . \quad (20)$$

where v and V_i are expressed in volts and J is the mechanical equivalent of heat.

Let the thermometer consist of a short length, say 5 mm.,

of platinum wire of radius ρ, placed along the axis Oz and in a region of the discharge where Z is independent of z.

For $\rho \leq r \leq R$ equation (10) still holds, q being given by equation (9). Integrating equation (10), and using the condition that $-2\pi rk\frac{\partial t'}{\partial r} = q'$ when $r=\rho$, we obtain

$$k\frac{\partial t'}{\partial r} + \frac{1}{2}qr = \frac{1}{2}\left(q\alpha^2 - \frac{q'}{\pi}\right).\frac{1}{r}, \quad . \quad . \quad . \quad (22)$$

where t' is now the temperature at a distance r from the axis Oz. Integrating equation (22) with the condition that $t'=t'_R$ when $r=R$, and solving for $t'-t'_R$, we obtain

$$t'-t'_R = \frac{-1+\sqrt{1+\frac{\alpha}{k_0}\left(\{R^2-r^2\}\frac{q}{2} + \left\{q\rho^2 - \frac{q'}{\pi}\right\}\log\frac{r}{R}\right)}}{\alpha} \quad . \quad . \quad . \quad (23)$$

To calculate t'_R, we have for $R \leq r \leq (R+\delta)$

$$\frac{1}{r}\frac{\partial}{\partial r}k'\frac{r\partial t}{\partial r} = 0.$$

Integrating this equation, and bringing in the condition that $-2\pi rk'\frac{\partial t}{\partial r} = \pi R^2 q + q'$ when $r=R$,

$$t'_R = t_0 + \frac{1}{2k'}\left\{R^2 q + \frac{q'}{\pi}\right\}\log_e\left(1+\frac{\delta}{R}\right). \quad . \quad (24)$$

Thus we have, on making the same approximations as in equation (15),

$$t' = t_0 + \frac{1}{2k'}\left\{R^2 q + \frac{q'}{\pi}\right\}\frac{\delta}{R} + \frac{1}{2k_0}\left\{(R^2-r^2)\frac{q}{2} + \left(\frac{q'}{\pi} - \rho^2 q\right)\log_e\frac{R}{r}\right\}. \quad (25)$$

Hence the difference between the rise in temperature as measured by the platinum-wire thermometer and the rise calculated from equation (16) will be

$$(t'-t)_{r=\rho} = \frac{q'\delta}{2\pi Rk'} + \frac{1}{2k_0}\left(\frac{q'}{\pi} - \rho^2 q\right)\log_e\frac{R}{\rho}. \quad . \quad (26)$$

Using the values of k', k_0, δ, R, and J given above, and assuming $\rho = \cdot0005$ cm., we obtain

$$(t'-t)_{r=\rho} = \cdot1105\{8\cdot01(v+V_i)I_+ - \cdot000000903\, iZ\}. \quad (27)$$

Assuming $\nu = 24$ volts, $V_i = 16$ volts, $I_+ = 100$ microamperes per centimetre, and $iZ = 500$ milliwatts per centimetre, *i. e.*, a large value of iZ, we find

$$(t'-t)_{r=\rho} = \cdot 1105\{32\cdot 04 - \cdot 00045\} = 3\cdot 54^\circ \text{ C.}$$

Hence the above calculation gives a correction to the readings of the platinum-wire thermometer which is of the required order of magnitude. Further, the above expression will give approximately the error of a thermometer wire which is placed a few millimetres from the axis Oz, since the rise in temperature produced in the discharge by the liberation of heat on the surface of the wire is very small near the inner surface of the discharge-tube. For example, in the case cited above, the rise in temperature at $r = 1$ cm. is $\cdot 183^\circ$ C., and at $r = 1\cdot 49$ cm. (*i. e.*, the inner surface of the discharge-tube) $\cdot 0074^\circ$ C.

Thus it would appear probable that the discrepancy between the rise of temperature in the discharge as measured by the platinum-wire thermometer and the rise calculated from the rate of dissipation of energy by means of equation (17) is due to the bombardment of the wire by positive ions and to the neutralization of the positive ions on the surface of the wire.

It might not be out of place to include in this paper the theory of the error associated with the two-probe wire method of measuring the electric force in a glow discharge. This is developed from I. Langmuir's theory of a cylindrical collector placed in the path of a glow discharge.

The method consists in measuring electrostatically the difference in potential, $\delta V'$, between two fine parallel insulated wires placed in the discharge at a known small distance, δz, apart. The electric force is then assumed to be $-\frac{\delta V'}{\delta z}$. The argument used to justify this is as follows: suppose one probe takes up a potential which is ν_0 volts below the actual potential of the surrounding discharge; then, since the other probe is similar and close to the first, it will also be at a potential ν_0 volts below the actual potential of the surrounding discharge. Thus both probes are in error by the same amount, and since only the difference in potential is measured, the error is eliminated.

The above is true in a uniform positive column, but is not true when the electric force is varying rapidly along the path of the discharge.

When an insulated probe is placed in the path of a dischage it takes up a potential which is negative with respect

to the surrounding discharge. The negative potential is sufficient to accelerate the positive ions to the probe and retard the electrons in order that the sum of the currents to the wire, due to the positive ions and electrons, may be zero. This is necessary, since the charge on the insulated probe must remain constant.

Let V be the actual potential at the point z on the axis of the discharge with respect to the cathode, and let V′ be the potential assumed by an insulated probe placed there. Then the potential assumed by the other probe at $z+\delta z$ will be

$$V' + \frac{\partial V'}{\partial z} \cdot \delta z + \frac{1}{2} \frac{\partial^2 V'}{\partial z^2} \cdot \delta z |^2 + \ldots ;$$

∴ measured electric force

$$= \frac{\delta V'}{\delta z} = \frac{\partial V'}{\partial z} + \frac{1}{2} \frac{\partial^2 V'}{\partial z^2} \cdot \delta z + \ldots$$

Let $$V - V' = \nu_0 ;$$

∴ measured electric force

$$= \frac{\delta V'}{\delta z} = \frac{\partial V}{\partial z} - \frac{\partial \nu_0}{\partial z} + \frac{1}{2} \frac{\partial^2 (V - \nu_0)}{\partial z^2} \cdot \delta z + \ldots$$

Assuming δz small, we have actual electric force

$$= \frac{\partial V}{\partial z} = \frac{\delta V'}{\delta z} - \frac{\partial \nu_0}{\partial z} . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (28)$$

The above argument assumes $\frac{\partial \nu_0}{\partial z}$ to be zero. A graph showing the variation of ν_0 with z is given by McCurdy (*loc. cit.*) for a striated discharge in mercury vapour.

When a cylindrical wire of radius ρ and length l is inserted into the path of a glow discharge and maintained at a potential ν volts below the actual potential of the surrounding discharge, the wire becomes surrounded by a positive ion sheath and a current, I_+, of positive ions flows to the wire. The problem was investigated by I. Langmuir, who obtained the following value for I_+*:

$$I_+ = \frac{2\sqrt{2}l}{9} \sqrt{\frac{e}{M}} \cdot \frac{\nu^{3/2}}{\rho \beta^2} . \quad . \quad . \quad . \quad . \quad (29)$$

The quantity β is a function of the ratio of the radius of the positive ion sheath to the radius of the wire. M is the mass and e the charge of a positive ion.

* Gen. El. Rev. xxvii. p. 452 (1924).

I_+ is independent of ν, the thickness of the positive ion sheath changing so that $\frac{\nu^{3/2}}{\beta^2}$ remains constant. Thus, if I_+ is measured at a potential which is known to be well below the actual potential of the discharge at the probe, then this will also be the value of the positive ion current to the wire when insulated.

In addition to the above current flowing to an insulated probe there is a current due to electrons which possess sufficient energy to penetrate the positive ion sheath against the retarding potential ν. According to Langmuir * the value of this current is

$$I_- = An_1 e \sqrt{\frac{e\overline{V}}{3\pi m}} \cdot e^{-\frac{3\nu}{2\overline{V}}}, \quad . \quad . \quad . \quad . \quad (30)$$

where

n_1 is the density of electrons surrounding the probe,

$\overline{V}$ their average energy in volts,

e and m their charge and mass, and

A the area of the probe.

Since the charge on the insulated probe remains constant, the potential assumed by it is given by

$$I_+ = I_-,$$

which is satisfied by a value of $\nu = \nu_0$ given by

$$\frac{2\sqrt{2}l}{9}\sqrt{\frac{e}{M}}\frac{\nu_0^{3/2}}{\rho\beta^2} = An_1 e\sqrt{\frac{e\overline{V}}{3\pi m}} \cdot e^{-\frac{3\nu_0}{2\overline{V}}}. \quad . \quad (31)$$

Differentiating equation (31) logarithmically, and solving for $\frac{\partial \nu_0}{\partial z}$, we obtain

$$\frac{3}{2}\left\{\frac{1}{\overline{V}} + \frac{1}{\nu_0}\right\}\frac{\partial \nu_0}{\partial z} = \frac{1}{n_1}\frac{\partial n_1}{\partial z} + \frac{1}{2\overline{V}}\left\{1 + \frac{3p}{V}\right\}\frac{\partial \overline{V}}{\partial z} + \frac{2}{\beta}\frac{\partial \beta}{\partial z}. \quad . \quad . \quad . \quad (32)$$

On differentiating equation (29) we find

$$\frac{\partial I_+}{\partial z} = I_+\left\{\frac{3}{2\nu_0}\frac{\partial \nu_0}{\partial z} - \frac{2}{\beta}\frac{\partial \beta}{\partial z}\right\}. \quad . \quad . \quad . \quad (33)$$

Now $\frac{\partial I_+}{\partial z}$ can be found by measuring the values of I_+ for the two probes when connected to a suitable low potential.

* *Loc. cit.* p. 449.

Solving equation (33) for $\frac{2}{\bar{\beta}}\frac{\partial\beta}{\partial z}$, and substituting in equation (32), we have

$$\frac{3}{2\bar{V}}\frac{\partial\nu_0}{\partial z}=\frac{1}{n_1}\frac{\partial n_1}{\partial z}+\frac{1}{2\bar{V}}\left\{1+\frac{3\nu_0}{\bar{V}}\right\}\frac{\partial\bar{V}}{\partial z}-\frac{1}{I_+}\frac{\partial I_+}{\partial z}. \quad (34)$$

By solving equation (30) for $\frac{\nu_0}{\bar{V}}$ we can eliminate $\frac{\nu_0}{\bar{V}}$ from equation (34) and obtain

$$\frac{3}{2\bar{V}}\frac{\partial\nu_0}{\partial z}=\frac{1}{n_1}\frac{\partial n_1}{\partial z}+\frac{1}{2\bar{V}}\left\{1+2\log\left(\frac{An_1e}{I_+}\sqrt{\frac{e\bar{V}}{3\pi m}}\right)\right\}\frac{\partial\bar{V}}{\partial z}$$
$$-\frac{1}{I_+}\frac{\partial I_+}{\partial z}. \quad . \quad . \quad . \quad (35)$$

In equation (35) there are still two unknowns, viz., n_1 and $\bar{V}$, and their z derivatives.

Let j be the current density in the discharge, then

$$j=en_1\bar{v}_1+en_2\bar{v}_2,$$

where $\bar{v}_1$ and $\bar{v}_2$ are the average velocities of the electrons and positive ions respectively parallel to Oz, and n_2 is the density of the positive ions.

Now n_2 is of the same order of magnitude as n_1, but $\bar{v}_2$ is very much smaller than $\bar{v}_1$ on account of the relatively large mass of the positive ions. Thus

$$j\doteqdot en_1\bar{v}_1. \quad . \quad . \quad . \quad . \quad . \quad . \quad (36)$$

Making use of the equation

$$\bar{V}e=\tfrac{1}{2}mv_1^2$$

we obtain

$$j=en_1\sqrt{\frac{2e\bar{V}}{m}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (37)$$

Hence we have from equation (35)

$$\frac{\partial\nu_0}{\partial z}=\frac{(2e\bar{V})^{3/2}}{3j\sqrt{m}}\frac{\partial n_1}{\partial z}+\frac{1}{3}\left\{1+2\log\left(\frac{Aj}{\sqrt{6\pi I_+}}\right)\right\}\frac{\partial\bar{V}}{\partial z}$$
$$-\frac{2\bar{V}}{3I_+}\frac{\partial I_+}{\partial z}. \quad . \quad . \quad . \quad (38)$$

This equation can be further reduced to known terms by using two equations given by Sir J. J. Thomson*, viz.:

$$j=n_1e\bar{v}_1+n_2e\bar{v}_2=e(n_1k_1+n_2k_2)\frac{\partial V}{\partial z}. \quad . \quad (39)$$

* Conduction of Electricity through Gases, 1st ed. p. 65.

and

$$\frac{\partial^2 V}{\partial z^2} = 4\pi(n_2 - n_1)e, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (40)$$

where k_1 and k_2 are the mobilities of the electrons and positive ions respectively.

Solving equations (38) and (39) for n_1, and differentiating, we have

$$\frac{\partial n_1}{\partial z} = \frac{1}{e(k_1 + k_2)} \left\{ \frac{-j}{\left| \frac{\partial V}{\partial z} \right|^2} \frac{\partial^2 V}{\partial z^2} - \frac{k_2}{4\pi} \frac{\partial^3 V}{\partial z^3} \right\}.$$

In accordance with the assumption in equation (39)

$$\overline{V}e = \tfrac{1}{2} m k_1^2 \left| \frac{\partial \overline{V}}{\partial z} \right|^2.$$

Hence, on eliminating $\frac{\partial n_1}{\partial z}$, $\overline{V}$, and $\frac{\partial \overline{V}}{\partial z}$ from equation (38) we obtain finally

$$\frac{\partial \nu_0}{\partial z} = \frac{m k_1^2}{3e} \left\{ 2 \log \left(\frac{A j}{I_+ \sqrt{6\pi}} \right) \cdot \frac{\partial^2 V}{\partial z^2} - \frac{1}{I_+} \frac{\partial V}{\partial z} \cdot \frac{\partial I_+}{\partial z} \right.$$

$$\left. - \frac{k_2}{4\pi j} \left| \frac{\partial \overline{V}}{\partial z} \right|^2 \frac{\partial^3 V}{\partial z^3} \right\} \frac{\partial V}{\partial z}. \quad . \quad . \quad (41)$$

In a uniform positive column, where $\frac{\partial V}{\partial z}$ is independent of z, $\frac{\partial^2 V}{\partial z^2}$, $\frac{\partial^3 V}{\partial z^3}$, and $\frac{\partial I_+}{\partial z}$ will be zero. Therefore in a uniform positive column $\frac{\partial \nu_0}{\partial z}$ is zero. This is confirmed by the close agreement between $V_1 + \int_A^B \frac{\delta V'}{\delta z} dz$ and V_2 for discharges of the type shown in fig. 4.

Summary.

1. It is suggested that the results of Messrs. Finch and Cowen's experiments upon gaseous combustion and ignition in electric discharges might be explained equally well upon the thermal theory.

2. Experiments made to test the accuracy of the two-probe method of determining the electric force in glow discharges are described. These lead to the conclusion that

above a pressure of about 2 mm. of mercury the method gives the electric force with an error of less than 1 per cent. in discharges passing through air or hydrogen. The appearance of a peculiar isolated striation found in air discharges is described, and it is suggested that it is similar to that observed in hydrogen discharges by Prof. J. Zeleny. Graphs are given showing the distributions along the axis of a discharge-tube of the rise in temperature obtained by means of a platinum-wire thermometer and of the corresponding rate of dissipation of energy. It is found that the observed rise in temperature is not exactly proportional to the rate of dissipation of energy.

3. The rise in temperature in the uniform positive column type of discharge is calculated from the rate of dissipation of energy in a manner similar to that employed by Prof. E. Warburg. This calculation shows that the rise in temperature should be, to a first approximation, proportional to the rate of dissipation of energy. The calculated values of the rise in temperature are found to be lower than the observed values. It is suggested that the difference between the rise in temperature indicated by the platinum-wire thermometer and the calculated value is due to the bombardment of the thermometer wire by positive ions, and to the subsequent neutralization of the positive ions by electrons on the surface of the wire. A calculation is given showing that this effect gives a correction which is of the correct order of magnitude. The theory of the error of the two-probe wire method of measuring the electric force in a glow discharge is developed from the work of I. Langmuir, and it is shown that the method gives the electric force exactly in a uniform positive column.

The above investigation was undertaken at the suggestion of Professor Taylor Jones, and the writer wishes to acknowledge his indebtedness to him for the interest and help which he gave throughout the course of the work. The experimental part of the investigation was carried out in the research laboratories of the Natural Philosophy Department of Glasgow University.

July 1930.

XVII. *The Transport Numbers of Hydrogen Chloride in Ethyl Alcohol.* *By* J. W. Woolcock, *B.A.*, Sir Harold Hartley, *F.R.S.*, *and* O. L. Hughes *.

THE values for these transport numbers given by previous workers are not very concordant. Lapworth and his coworkers †, using the Hittorf method, have obtained results for η_H which do not agree very well amongst themselves, but indicate a value lying between 0·65 and 0·70. Harned and Fleysher ‡ have employed the MacInnes method § of comparing the electromotive forces of cells of the types

$$\mathrm{Ag, AgCl}\left|\begin{matrix}\mathrm{HCl}\\ m_1\end{matrix}\right|\mathrm{H_2-H_2}\left|\begin{matrix}\mathrm{HCl}\\ m_2\end{matrix}\right|\mathrm{AgCl, Ag}\ldots\ldots\mathrm{E_1}, \quad (1)$$

$$\mathrm{Ag, AgCl}\left|\begin{matrix}\mathrm{HCl} & \mathrm{HCl}\\ m_1 & m_2\end{matrix}\right|\mathrm{AgCl, Ag}\ldots\ldots\mathrm{E_2}. \quad (2)$$

When the values of E_1 and E_2 are plotted against the logarithm of the activity, the transport number of the hydrogen ion at any concentration is given by the ratio of the slopes of the curves.

This method has two great advantages :

(1) The values of η_H are determined from a smooth curve, and not from isolated points, and they refer to definite concentrations.

(2) The value of η_H at infinite dilution (which is of the greatest importance for assigning mobilities to individual ions) can be determined, for this is the value of η_H when the E_1 and E_2 curves have become straight lines as the solution becomes more dilute.

Harned and Fleysher calculated values for η_H by assigning to the E_2 curve the form

$$E_2 = B + C \log A + D(\log A)^2,$$

where A is the activity of the solution $\times 10^4$, which gives

$$\eta_H = K + L \log A.$$

* Communicated by the Authors.
† Lapworth, Journ. Chem. Soc. cxvii. p. 857 (1915).
‡ Harned and Fleysher, Journ. Amer. Chem. Soc. xlvii. p. 92 (1925).
§ MacInnes and Beattie, *Ibid.* xlii. p. 1117 (1920).

This method, while useful in stronger solutions, does not allow for the straightening of the E_2 curve in dilute solution, and consequently yields varying values for η_H over the range of concentration where there is good reason for thinking that the transport number is constant. At 0·001 *m* they find $\eta_H = 0{\cdot}753$ from this formula, but at this dilution the E_2 curve has become a straight line, and the slope of the line indicates that 0·73 is the probable value for η_H. It has already been pointed out * that the results obtained by Harned and Fleysher are probably in error owing to the

Values of the electromotive forces of cells with and without transport, and the transport numbers calculated therefrom.

m_1.	m_2.	E_1.	E_2.	η_H.
1·2411	0·1364	0·0685	0·0524	0·765
1·0000	0·1000	0·0725	0·0542	0·748
0·8388	0·08238	0·0745	0·0540	0·725
0·7457	0·07319	0·0720	0·0535	0·743
0·4753	0·04171	0·0795	0·0581	0·731
0·1364	0·01062	0·0895	0·0644	0·720
0·08238	0·008224	0·0822	0·0579	0·704
0·06239	0·004111	0·1000	0·0705	0·705
0·02492	0·002461	0·0895	0·0646	0·722
1·2411	0·004111	0·1950	0·1418	0·727
1·0000	0·01035	0·1530	0·1129	0·738

presence of traces of water in their alcohol; so that too much reliance must not be placed upon their values for η_H.

Drucker and Schingnitz † have recalculated the results of Harned and Fleysher, obtaining transport numbers which are the mean values between the concentrations *m* and 0·01 *m*. The latter value was rather an unfortunate choice, as it lies somewhat off the E_1 curve, thereby vitiating their calculations.

The electromotive force of the cell (2) has been redetermined at 25°C., using a cell with a flowing junction of the type described by Nonhebel and Hartley ‡, and the results obtained are given in the table. The experimental details

* Woolcock and Hartley, Phil. Mag. v. p. 1133 (1928).
† Drucker and Schingnitz, *Zeit. Phys. Chem.* cxxii. p. 149 (1926).
‡ Nonhebel and Hartley, Phil. Mag. i. p. 729 (1925).

were identical with those described by Woolcock and Hartley (*loc. cit.*), and the values of E_1 are taken from their results.

Unfortunately, the results do not permit of treatment by MacInnes' method. The values of η_H have a definite trend, decreasing with dilution, and tending to a value of about 0·71. The change with dilution is in an opposite direction to that found by Harned and Fleysher, but in the same direction as that observed for η_H in methyl alcohol by Nonhebel and Hartley (*loc. cit.*). The irregularities in the transport number are probably due to the very large effect of small traces of water, unevenly distributed between the two solutions, on the electromotive force of the cells.

Murray-Rust and Hartley * have found a value $\lambda_0 = 83{\cdot}8$ for hydrogen chloride in ethyl alcohol at 25° C., so that, taking $\eta_{0\,H} = 0{\cdot}71$, we have

$$l_{0\,H} = 59{\cdot}5 \quad \text{and} \quad l_{0\,Cl} = 24{\cdot}3.$$

Using these mobilities as standards, it is possible to calculate ionic mobilities for other ions from conductivity measurements (not yet published) made in this laboratory by Barak, Copley, and Murray-Rust, and also the transport numbers for various salts at infinite dilution. It is interesting to compare these with certain directly determined values.

Silver Nitrate.

$$l_{0\,Ag} = 17{\cdot}9 \; : \; l_{0\,NO_3} = 28{\cdot}1 \; : \; \eta_{0\,Ag} = 0{\cdot}39.$$

Krumreich † found $\eta_{Ag} = 0{\cdot}39$ from electromotive force measurements, and $\eta_{Ag} = 0{\cdot}41$ from direct transport measurements; but since he worked in 0·01 N solutions, and at 40° C., his results are not strictly comparable with ours.

Gibbons and Getman ‡ found $\eta_{Ag} = 0{\cdot}397$ by the Hittorf method in 0·05 *m* solution at 25° C.

Potassium Iodide.

$$l_{0\,K} = 22{\cdot}0 \; : \; l_{0\,I} = 28{\cdot}7 \; : \; \eta_{0\,I} = 0{\cdot}56.$$

Robson, in a preliminary investigation made in this laboratory, found values for η_I of 0·52, 0·53, 0·54 at 25° C., with a preference for the higher values, using the Hittorf method at concentrations of about 0·01 N. Isaacs and

* Murray-Rust and Hartley, Proc. Roy. Soc. A, cxxvi. p. 84 (1929).

† Krumreich, *Zeit. für Elect.* xxii. p. 446 (1916).

‡ Gibbons and Getman, Journ. Amer. Chem. Soc. xxxvi. p. 1630 (1914).

Partington* have measured the electromotive force of the cell

$$\mathrm{Ag,\ AgI}\left|\begin{matrix}\mathrm{KI} & \mathrm{KI}\\ c_1 & c_2\end{matrix}\right|\mathrm{AgI,\ Ag},$$

and have calculated transport numbers from the equation

$$\mathrm{E}=2\,[1-\eta_{\mathrm{I}}]\,\frac{\mathrm{RT}}{\mathrm{F}}\,\log_e\frac{c_1\lambda_1}{c_2\lambda_2}.$$

They obtain a series of concordant values with a mean of $\eta_{\mathrm{I}}=0{\cdot}55$, but their use of this form of the Nernst equation assumes that the change in activity coefficient is proportional to the change in equivalent conductivity.

Sodium Iodide.

$$l_{0\,\mathrm{Na}}=18{\cdot}7\ :\ l_{0\,\mathrm{I}}=28{\cdot}7\ :\ \eta_{\mathrm{I}}=0{\cdot}605.$$

Partington and Simpson† measured the electromotive forces of cells with and without transport and found the value $\eta_{\mathrm{I}}=0{\cdot}58$ at 0·01 N.

The values found directly for these three salts are in fairly good agreement with the transport numbers given above for hydrogen chloride.

Summary.

Measurements have been made of the E.M.F.s of cells containing hydrogen chloride in ethyl alcohol, with and without liquid junctions, and the most probable value for the transport number of the hydrogen ion at infinite dilution has been found to be $0{\cdot}71\pm0{\cdot}01$.

Physical Chemistry Laboratory,
Balliol College and Trinity College, Oxford.

XVIII. *Comments on the Paper by Shilling and Laxton on "The Effect of Temperature on the Viscosity of Air."*

To the Editors of the Philosophical Magazine.

GENTLEMEN,—

IN the October (1930) issue of your Magazine Dr. W. G. Shilling and Mr. A. E. Laxton have described their measurements of the temperature variation of the viscosity

* Isaacs and Partington, Trans. Faraday Soc. xxv. p. 53 (1930).
† Partington and Simpson, Trans. Faraday Soc. xxvi. p. 625 (1930).

of air. These measurements were undertaken as a consequence of my criticisms * of the results obtained previously by Williams †, and were virtually a repetition of Williams's work. His method was modified so as to investigate the suggestions I had made as possibly accounting for the exceptional conclusion that Williams reached, namely, that Sutherland's law was not even approximately true over the temperature range 0°–250° C., and only became valid (with a value of the constant much higher than that usually accepted) above that range.

Shilling and Laxton claim that their investigations provide a complete vindication against all my criticisms of the experimental method used by Williams, yet at the same time they find their results disagreeing widely from his and failing to confirm his surprising conclusions. It seems to me that the real importance of their work lies in helping to show that Williams's results are in error, and that the exceptional effect which provoked my criticisms is non-existent, rather than in their partial proof that my suggestions as to the possible causes of the discrepancies happen to have been incorrect. The position now is that in three different investigations of the matter made subsequently to the publication of Williams's results, namely by Edwards ‡, by Braune, Basch and Wentzel §, and now by Shilling and Laxton themselves, no confirmation is to be found of a breakdown of Sutherland's law in the manner indicated by Williams. For even Braune, Basch and Wentzel, to whom Shilling and Laxton refer (p. 733), although apparently not unwilling to accept Williams's conclusions, indicate that in their own range of measurements no real departure from Sutherland's formula can be affirmed with certainty.

Although admitting the non-existence of the sudden departure from Sutherland's law which Williams claims to have discovered, Shilling and Laxton still affirm (p. 733) that there is a small but steady deviation from this law over the temperature range observed, as indicated by Edwards's work as well as their own. Even this appears to me to be a deduction not permissible from the observations. The authors themselves give no estimate of their

* Proc. Roy. Soc. A, cxi. p. 219 (1926).
† *Loc. cit.* p. 141.
‡ Proc. Roy. Soc. A, cxvii. p. 245 (1927).
§ *Zeit. Phys. Chem.* cxxxvii. p. 176 (1928).

probable errors, but an examination of the individual results on p. 731 provides internal evidence that these may be more than 3 per cent. in error. For example, the four results in order in the third column of Table VII. from 182° to 189° C., actually show a *diminution* of 58 parts in about 2500, whereas the actual *increase* of viscosity for 7° rise in temperature amounts in this region to about 26 parts per 2500. Together these are equivalent to about 84 parts in 2500, or 3·4 per cent. In this connexion it should be noted that the two points corresponding to the highest temperatures on the deviation diagram (fig. 2, p. 732) do in fact represent *single* observations, and are presumably therefore liable to be something like 30 parts per thousand in error. Yet their deviations from Sutherland's law amount to only about 8 parts per thousand. In view of the comparatively large experimental errors these small deviations are accordingly without significance; nevertheless, the upward trend indicated and stressed depends chiefly upon them. On the same grounds it is reasonable to entertain serious doubts whether any of the deviations represented are large enough, in relation to the experimental accuracy attained, to justify any conclusion other than that of the validity of Sutherland's law over the whole temperature range covered.

In the case of Edwards's results, included in the same diagram, the deviations have been exaggerated by failure to find the best fit for Sutherland's equation, for, if this were done, the points could not all lie on one side of the zero line. Even so, the deviations represented are just of the same order as Edwards's * own estimates of the probable experimental errors, and imagination seems to be the basis of the upward trend suggested by the dotted line.

Yours faithfully,

A. O. Rankine.

Imperial College of Science and Technology,
October 17, 1930.

* *Loc. cit.* pp. 254-256.

XIX. *Notices respecting New Books.*

The Structure of Line Spectra. By L. Pauling, Ph.D., and S. Goudsmit, Ph.D. [Pp. x + 263, with 68 figures.] (New York: McGraw-Hill Book Co. London: McGraw-Hill Publishing Co. 1930. Price $3.50.)

THE investigation of the structure of line-spectra, both on the theoretical and practical sides, has been developed during recent years with many ramifications and in great detail. It is not easy for the worker in the field of spectroscopy to keep count of the flood of papers which are published, and the variety of nomenclatures which have been or are in use is a further complication. The volume under review gives an excellent summary of present knowledge. It has been written as a textbook for those who are entirely unacquainted with the subject and for those who are familiar with the empirical side of spectra, but not with their interpretation.

The authors give an exposition of the Bohr-Hund theory. The treatment is based on a vector model which has the advantage of being easily visualised and being less abstract than the methods of quantum mechanics. Wherever necessary, the slight modifications introduced into the results by quantum mechanics are quoted and used. A minimum amount of mathematics is employed, and results which are based upon more complicated mathematics are quoted without proof.

The authors have not included an account of the structure of all line spectra which have been analysed, but typical spectra are discussed for illustrative purposes. They have concerned themselves mainly with the general principles involved. The material is well-arranged and the discussion throughout is characterised by great lucidity and simplicity.

Beginning with a general account of atomic theories and atomic models, a discussion of the stationary states of the hydrogen atom follows. Alkali-like atoms are then dealt with. The conception of the spinning electron is introduced and the vector model is explained. This is applied to atoms with one- and two-valency electrons and then to many-electron atoms. Einstein transition probabilities are then introduced as a prelude to intensity considerations. The Pauli exclusion principle is next explained and the structure of atoms in conjunction with the periodic system of the elements is discussed. Chapters on X-ray spectra, hyperfine structure, and magnetic phenomena complete the volume.

There are naturally many omissions which could be mentioned. Encyclopædic treatment is not possible in a volume of this size. The authors have presented a well-balanced account for which spectroscopic workers will be grateful.

The Conductivity of Solutions and the Modern Dissociation Theory. By Cecil W. Davies, M.Sc., A.I.C. [Pp. viii+204.] (Chapman & Hall, 1930. Price 15*s.*)

The author summarises the development of modern ideas about the nature of electrolytes in solution, and reviews the present situation in this field. The first section of the book is devoted to a historical and theoretical introduction which, although comprehensive, is commendably concise. The second section—"Methods and Results"—includes the principal experimental details of conductivity measurement, and discusses the results actually obtained in a wide range of solutes, solvents, and concentrations. Particular attention is paid to the application of the Onsager equation, but the empirical formulæ of other workers receive due mention. Although this will probably prove the most valuable part of the book from the point of view of the student, the last section—"Some Applications and Consequences"—dealing with such questions as Bronsted's theory of acids and bases, solvolysis, and complex ions, is undoubtedly the most fascinating in the possibilities which it suggests. It is interesting to note that the last chapter deals with the "degree of dissociation" of electrolytes, and it is gratifying to find throughout the book a proper estimation of the value of the Complete Dissociation Theory in explaining the behaviour of such solutions as most chemists have to work with. Now that conductivity workers are beginning to attack more general cases and more concentrated solutions there is no reason why their conclusions should not have some measure of that influence which the theory of Arrhenius had upon the subject as a whole. Mr. Davies is prepared to discuss these questions, and presents the case with unbiassed judgment. He has written an eminently readable book, and has undoubtedly achieved the objects outlined in his preface.

Intermediate Mechanics. By D. Humphrey, B.A., B.Sc. [Pp. xi+382, with diagrams.] (London: Longmans, Green & Co, 1930. Price 10*s.* 6*d.*)

Designed to meet the requirements of students working for Intermediate B.Sc., various Higher Certificate and University Scholarship Examinations, this book deals with harder applications of the principles of the subject than is customary in a school text-book. It is not intended for beginners, but will be found suitable for those who have some previous acquaintance with the subject and an elementary knowledge of the calculus. The methods of the calculus are used throughout. It is advantageous to introduce such methods at as early a stage as possible; the treatment

is simplified and the fundamental ideas of the calculus are illustrated.

The subject-matter includes a section on motion under a central force, a chapter on motion of a rigid body about a fixed axis, and one on the motion of a rigid body in two dimensions. The general principles are illustrated by a large number of fully-worked examples. and many additional examples to be worked by the student are given. The volume can be thoroughly commended as a text-book for examinations of the standard mentioned above.

A second volume, dealing on similar lines, on statics and hydrostatics is promised.

Molecular Magnetism. By W. Peddie. (Edward Arnold & Co., London. Price 10*s.* net.)

Modern Physics. By Theodor Wulf. Translated from the German by C. J. Smith, Ph.D., M.Sc., A.R.C.S. (Methuen & Co., Ltd., 36 Essex Street, London, W.C. Price 35*s.* net.)

Physique moderne. By G. Castelfranchi. Translated from the Italian by M. A. Q. de Lanascol. (Librairie Scientifique Albert Blanchard, Place de la Sorbonne, Paris. Price 70 fr.)

The advances in the study of magnetism during the last quarter of a century have been as notable as in other branches of physics ; numerous text-books have appeared on X-rays, quantum theory, and atomic structure, but, with the exception of Dr. Stoner's books, no English work has been published giving the results of recent theoretical and experimental researches in magnetism. Prof. Peddie's 'Molecular Magnetism' supplies this want and gives a clear and helpful account of the researches of Weber, Curie, Langevin, and Weiss, and the more recent work of Honda, Webster, Forrest, and others. It will serve as a useful guide to those who may be stimulated to undertake research in this subject.

The two text-books on modern physics attempt to describe some of the more important physical developments. The aim of the former is to give a coordinated account of the fundamental results of physics for the use of the general reader ; mathematical formulæ are largely dispensed with in the interest of students whose equipment is not very extensive, but this scarcely justifies the doubtful derivation of potential "without using the calculus." The most recent topics selected for discussion include electromagnetic waves, ions and electrons, radioactivity, and the structure of the atom.

'Physique moderne' covers a much wider field, and presents a comprehensive and adequate survey of present-day physics, the kinetic theory of gases, the electron, X-rays, and crystal structure and radioactivity ; black body radiation and Planck's constant, the photoelectric and Compton effects.

The last three chapters deal with magnetism and quanta, wave mechanics, and the Fermi and Bose statistics. The experiments of Gerlach and Stern and of Davisson and Germer are clearly described and illustrated. This book is a model of clear exposition, and, from its high merits, has received the commendation of many of the leading physicists of the day.

The Electromagnetic Field. By Max Mason and Warren Weaver. (The University of Chicago Press, Chicago, Illinois. Price 27*s.* net.)

In this work, an introduction to the mathematical field theory of electrodynamics, written in an attempt to keep clear the relation between the mathematical mechanism and physical reality, the authors have succeeded in giving a clear and attractive presentation of the classical electrodynamic theory. The fundamental equations of electrostatics and magnetostatics are developed along similar lines with the introduction of the concept of continuous densities and the determination of the potential of concentric spherical shells, ellipsoid, and circular disk. Two special cases of the electrostatic problem for dielectrics are considered—that of a dielectric sphere in a uniform external field and of a uniformly polarized dielectric. The expressions for the scalar and vector potentials are applied to the problems of a uniformly moving point charge, the Abraham and Lorentz electrons, and the oscillating dipole. Vector methods are exclusively used: the operations of vector analysis, differential and integral, the potential and wave equations, and the general theory of curvilinear coordinates are set out in a mathematical appendix. A number of problems are given at the end of the chapters and an index provided of the numerous formulæ occurring in the text. Three of the names associated with the development of electrodynamic theory should be Riemann, Kirchhoff, and Helmholtz.

Sound Waves and their Uses. By Alexander Wood, M.A., D.Sc. [Pp. x+152, with 141 figures.] (London: Blackie & Son, Ltd., 1930. Price 7*s.* 6*d.* net.)

The Christmas Lectures delived before a juvenile auditory at the Royal Institution have been responsible for the publication of a number of excellent popular books on scientific subjects. Amongst such the present volume, adapted from the Christmas Lectures, 1928, may be included. The treatment of the subject has been to some extent modified, so that the volume will appeal to the elders rather than to the children. It provides an excellent introduction to recent developments of the subject, which cover a very extensive field. Amongst the subjects which are dealt with are the zones of silence in the transmission of sound, methods of echo-sounding, methods

of fixing the direction of a source of sound, ultra-sonic waves, the acoustics and acoustic correction of buildings, and methods of recording and reproducing sounds, including the gramophone and talking films. The analysis of musical notes and the choice of a musical scale are clearly discussed, and some account is given of recent work on the analysis of speech, and particularly of vowel sounds.

The volume is well printed and illustrated with numerous figures and diagrams.

Intermediate Dynamics and Properties of Matter. By R. A. Houstoun, M.A., D.Sc. [Pp. ix+139, with 138 figures.] (London : Longmans, Green & Co., 1929. Price 3*s.* 6*d.*)

Intermediate Physics. By R. A. Houstoun, M.A., D.Sc. [Pp. xviii+638, with 588 figures.] (London : Longmans, Green & Co., 1930. Price 10*s.* 6*d.*)

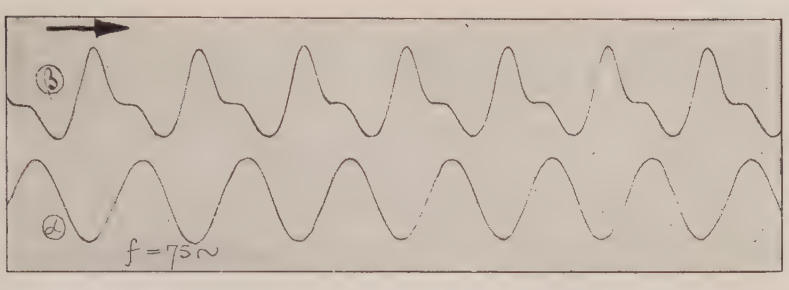
The second of the above books is divided into five parts, dealing respectively with dynamics and properties of matter, heat, sound, light, magnetism and electricity. With the exception of the section dealing with sound, these sections have been previously published separately. The first of the above books forms the initial section of the 'Intermediate Physics,' but was the last of the separate sections to be published.

The standard of the work is approximately that of the intermediate science degrees of British Universities, but the book is also well adapted for use as a text-book in schools. The treatment is very clear, the book is abundantly illustrated with excellent diagrams, and practical applications of theoretical matter are referred to whenever possible. A collection of questions from various public examinations enhances the value of the book as a text-book. The subject-matter is thoroughly up-to-date and has been selected with evident care by the author, who is an experienced teacher. The volume can be strongly recommended as a text-book.

The sections dealing with dynamics and properties of matter and with sound are new. The former covers a wide range, including kinematics, statics, dynamics, hydrostatics, and the general properties of matter. An interesting chapter is devoted to various types of pumps. The section on sound is brief : one chapter deals with wave motion, a second with the production and propagation of sound, a third with the measurement of frequency and the musical scale, and the last with vibrations of strings and columns of air with application to the phonograph and gramophone.

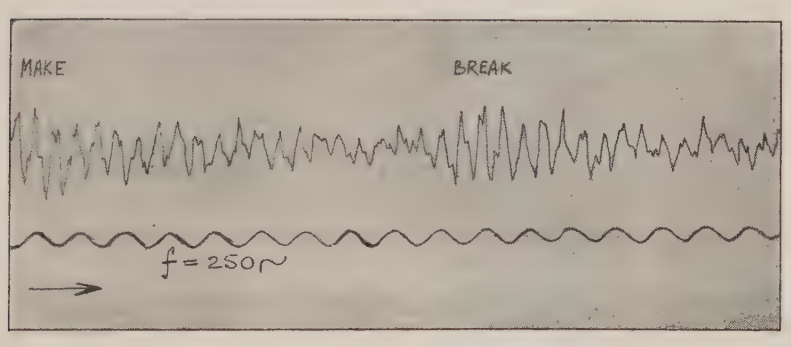
[*The Editors do not hold themselves responsible for the views expressed by their correspondents.*]

Fig. 8.



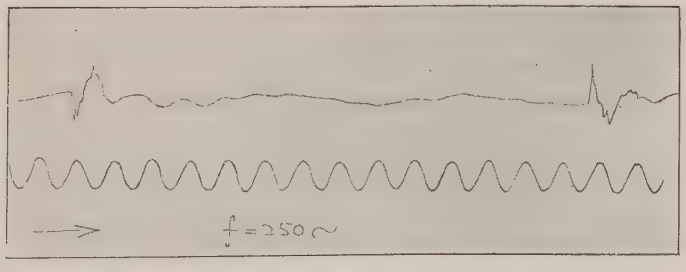
Record of acoustic output from reed-driven conical diaphragm ("Lion") when the amplitude of the reed exceeds the limit of linearity. Curve (α)=input to windings; curve (β)=acoustic output from diaphragm.

Fig. 18.



Wave-form of acoustic output from reed-driven aluminium disk 10 cm. rad. 0·165 cm. thick, corresponding to square wave input.

Fig. 19.



Wave-form of acoustic output from reed-driven diaphragm ("Lion") 21 cm. rad. with square wave input.

Fig. 12 *a*.

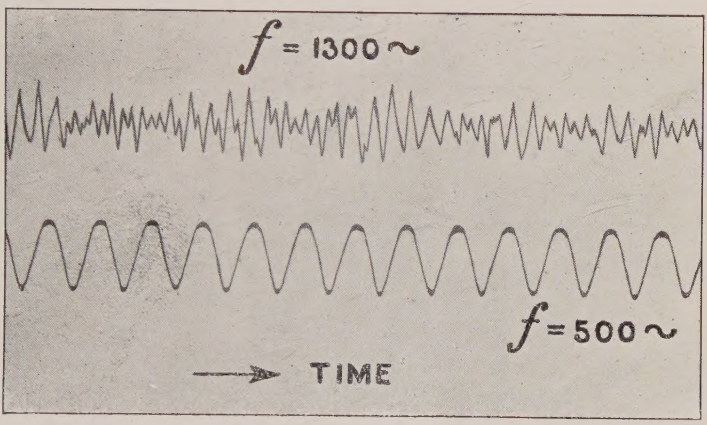
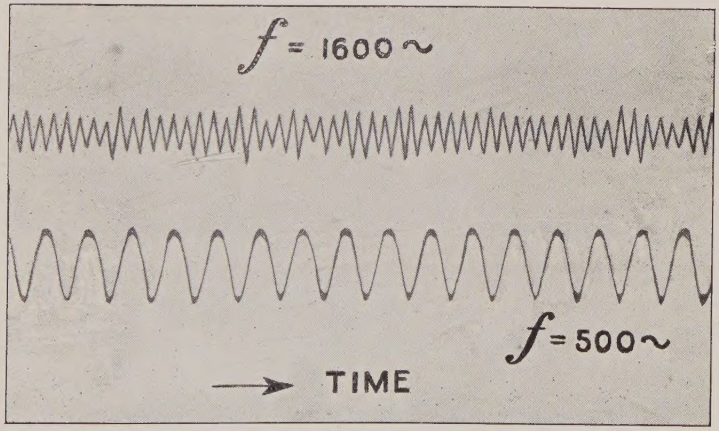
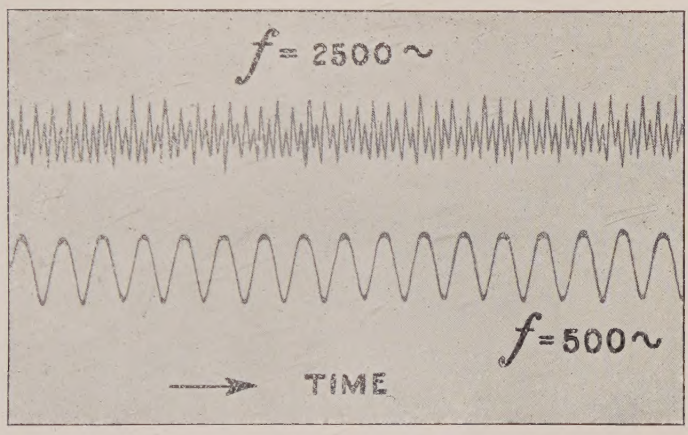


Fig. 12 *b*.



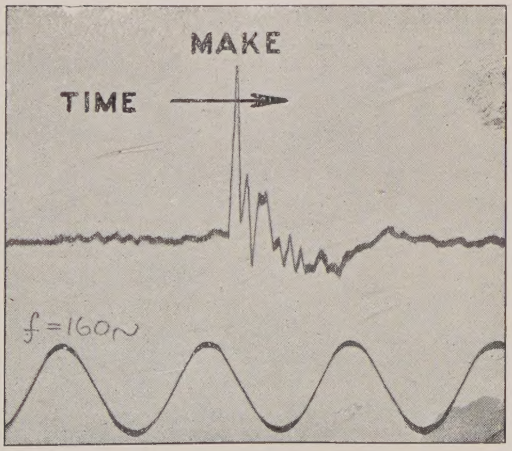
Records illustrating frequencies of oscillation of loud speaker coil when excited by violin bow.

Fig. 12c.



Record illustrating frequencies of oscillation of loud speaker coil when excited by violin bow.

Fig. 20.



Wave-form of acoustic output from coil-driven diaphragm 12 cm. rad. with square wave input.